TSSR

Trinity Student Scientific Review

Volume III

TSSR

Trinity Student Scientific Review

Volume III

Published by the Trinity Student Scientific Review
c/o Faculty of Engineering, Mathematics and Science,
Trinity College, Dublin 2
Republic of Ireland.

Print managed by Custodian.

Cover by Maria Cordero.

Photography by Bogdan Hrechka.

TSSR logo created by Ger Walsh at Magnet Design in 2014.

This journal is also available online at trinityssr.com

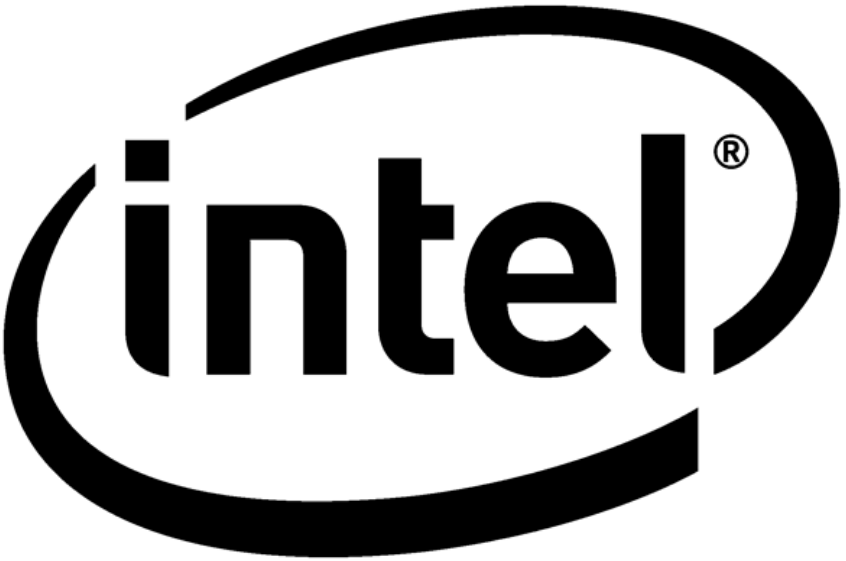
Inquiries at trinssr@tcd.ie

TS
SR

The Trinity Student Scientific Review is kindly sponsored by

The Faculty of Engineering, Mathematics and Science
Trinity College Dublin

Sponsorship



Sponsoring Scientific Research at Trinity College Dublin

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

With Thanks To:
The Faculty of Engineering Mathematics and Science
Trinity College Dublin

Editorial Team

General Manager	Sarah Jennings
Publications Manager	Maria Cordero
Life Sciences Editor	Jack Schofield
Assistant Life Sciences Editor	Kat Hughes
Assistant Life Sciences Editor	Conor McLoughlin
Chemistry Editor	Stephen Byrne
Natural Sciences Editor	Cian White
Physics Editor	Jeffrey Lyons

Prize-Winning Reviews

Best Overall Review

Resurrection Plants: The Evolution of Desiccation Tolerance and Potential Crop Improvement
Andrew Neill

Best Freshman Review

The Snowball Earth Hypothesis
Ellen Mullarkey

Best Life Sciences Review

Immunotherapy: At the Front Line in the Fight Against Cancer; CTLA4 and PD-1 Checkpoint Inhibitors
Eoin Connor

Best Chemistry Review

Epigenetic Cytosine Modifications and their Detection
Mark Berney

Best Physics Review

Designing Artificial Photosynthetic Systems Using Nanomaterials for Energy Production and Storage
Aoife Gregg

Acknowledgements

Natural Sciences Editor	Cian White
Natural Sciences Academic Advisor	Dr. Matthew Saunders
Natural Sciences Peer Reviewers	Dr. Trevor Hodkinson
	Maureen Williams
	Darren O'Connell
	Andrew Mooney
	Sabrina Renken
Life Sciences Editor	Jack Schofield
Assistant Life Sciences Editors	Kat Hughes
	Conor McLoughlin
Life Sciences Academic Advisor	Dr. Aisling Dunne
Life Sciences Peer Reviewers	Killian Hanlon
	Nicole Campbell
	Dinorath Olvera Ramos
	Jessica Walls
	Elena Stavenschi
	Proinnsias Fox
	Karen Slattery
	Natalie Adlesic
Chemistry Editor	Stephen Byrne
Chemistry Academic Advisor	Dr. Mike Southern
Chemistry Peer Reviewers	Dr. Maria Daniela Angione
	Dr. Michael Lyons
	Dr. Niall McEvoy
	Dr. Joanna McGouran
	Maria O'Brien
	Dr. Wolfgang Schmitt
Physics Editor	Jeffrey Lyons
Physics Academic Advisor	Prof. Igor Shvets
Physics Peer Reviewers	Subhayan Roychoudhury
	John Boland
	Stephen O' Brien
	Adam Kelly

Contents

Natural Sciences

Contents

Life Sciences

Contents

Contents

Chemistry

Contents

Physics

Welcome

Dear Reader,

Today the flow between ground-breaking science to our mundane daily lives has never been more strong and constant. We are incredibly lucky to have at our disposal the life's work of scientists across the globe, making use of Nobel prize winning LEDs in our regular lives and boggling at the potential of quantum computers in our futures. It remains, however, imperative to defend the value of science and justify its existence.

Much of this comes from enjoying the wonders of science- from the beauty of space to the hope of new disease cures. But as science students we also appreciate the more technical aspects of science. The literature reviews featured in this journal are at the edge of scientific research and serve as enlightening and exceptional windows to the research they draw from. They take a critical lens to the past and future of their fields, fields we hope the wonderful authors who submitted to this volume will enter in their careers and drive forward.

Founded in 2015 by an inspiring group of undergraduates, the TSSR became the first peer-reviewed undergraduate scientific review journal of its kind in Ireland. This journal represents the pinnacle of student undergraduate scientific research. It demonstrates the work done by students in their studies and beyond, and fosters a crucial understanding of scientific research amongst students. To the founding team I extend my gratitude. This wonderful publication would not exist without their initiative and hard work.

TSSR looks to be a bridge between the lecture hall and the world of research, both of which Trinity exceed in. Much is spoken of the past of Trinity Science, successful alumni and Nobel Laureates from Ernest Walton to William Campbell. Their successes serve as inspirations to every student in this university, but so too do the outstanding lecturers and researchers of the college. If the history of Trinity Science is well documented, and the present proves itself through the research it produces, then the future can be seen within this volume. Containing truly excellent multidisciplinary work, this year's edition speaks to the talent, passion, and curiosity of the students of Trinity. I am confident, dear reader, you will feel as hopeful on reading their work as I do.

Volume III has been a labour of many and they all deserve extended thanks.

To every student who submitted in record numbers this year I say thank you. The work submitted this year left us often in awe, and to do so on top of a gruelling degree is a testament to every one of you. I hope you gained something from the process regardless of publication and will look back on the experience with fondness.

To all those members of academic staff who gave their time, expertise and advice I extend my thanks and that of my team. As students we rely on the knowledge of experts and you were happy to provide it in spades.

To our sponsors- Intel and The Science Gallery- I am sincerely grateful. To Prof. Vinny Cahill, Dean of the Faculty of Engineering, Mathematics and Science, and everyone in the Faculty office I simply cannot thank you enough. Your support, kind words, and advice have helped deliver this journal to print. We simply couldn't have done it without you.

To last year's team, and in particular Alison, I cannot make my gratitude clear enough. This volume had the strongest of foundations to build upon from last year's edition and the best of supports along the way.

And finally, to my wonderful Editorial Board and Committee: Cian White, Jeffrey Lyons, Stephen Byrne, Jack Schofield, Kat Hughes, Conor McLoughlin and Maria Cordero. You have been a wonderful network on which to rely, have produced excellent work, and most importantly have been a first-rate group to know for a year. Thank you for being my teammates and my friends.

Finally passing from us to you, I present Volume III. I wish you pleasant reading.

Sarah Jennings

General Manager
Trinity Student Scientific Review 2017

Welcome

Dear Reader,

The TSSR endeavours to give students the opportunity to experience the process of academic writing. It aims to acknowledge and applaud the ability of these students to questions the world around them. Over the last 20 years, technology has revolutionised our relationship with scientific research. With the growth of the internet, we are now able to access information more quickly and easily than ever before. As such it's a testament to the published students' ability to gather and process the wealth of research done on their chosen topic and to refine it into the polished reviews in this volume.

As research has become more far-reaching it is now more important than ever to keep this research accessible. Science is central to finding the solutions to contemporary issues; climate change, the energy crisis, declining biodiversity. These are just a few examples of issues dealt with by some of the reviews published in this volume. It is increasingly important to ensure that scientific research is not held back or silenced in today's world.

Unrestricted access to research is essential in ensuring that we can enjoy evidence-based policy and reap the benefits of research in almost every aspect of our lives. Efforts to keep scientific discoveries like human genome mapping in the public domain underpin the spirit of science. The crucial role that scientific discoveries play in our lives is not one that can be obscured. In the face of efforts to discredit scientific findings or to restrict researchers we must continue to uphold scientific integrity. In essence, science strives only to find answers. In a world in need of these answers, we must continue to appreciate and share the scientific research that we as a student body hold so dear.

This journal exemplifies the open access to research that universities and the population as a whole deserve access to. In order to maintain its high calibre of work and also remain a journal that is free to the public, we have many people to thank.

First of all we are eternally appreciative of the work that all of the students who submitted reviews have done. Every year the editorial team expresses their amazement at the quality of the work submitted and every year that amazement is genuine. Secondly this journal could not pride itself on its high standard without the selfless work of the academic advisors and peer reviewers. These academics give up their time not only to ensure that the publication is exemplary but also that

the students who submit can experience the editorial process first-hand. Lastly I'd like to acknowledge my fellow members of the editorial board whose work over the past few months has been tremendous and without which I would have no publication to manage. It's been a pleasure working with you.

Maria Cordero

Publications Manager

Trinity Student Scientific Review 2017

Natural Sciences

TS
SR

Letter from the Editor

In light of the current political climate the importance of scientific communication has never been greater. It is vital that endeavours that seek not only to be bastions of fact but also educate people to think critically are supported and encouraged. It therefore gives me much pleasure to see the enthusiasm that both students and academic advisers have for the Trinity Student Scientific Review, illustrated by volume and quality of submissions that the TSSR received this year.

The third publication of the Trinity Student Scientific Review marks the renaming of this section as the Natural Sciences. This shifts the TSSR to be more in line with the how the schools within Trinity are structured and suitably encompasses the broad range of disciplines that study the natural world. The diversity of reviews submitted this year mirrors the broad scope of the section, from the evolution of plants and the potential for crop improvement, past and present effects of climate change on both biodiversity and geology, to conservation of mega herbivores using novel and controversial techniques. I hope you will find the papers to be both fascinating and informative.

It remains for me to thank the academic adviser Dr. Matthew Saunders for guidance and support. This extends to all who reviewed papers for this section; Dr Trevor Hodkinson, Darren O'Connell, Andrew Mooney, Maureen Williams and Sabrina Renken. It has been a pleasure and an educational experience to work alongside you.

Finally, I'd like to thank all the authors who contributed their work. It is daunting submitting a review that will be critically analysed and I do not underestimate the commitment it takes to produce a paper. It is my hope that each author improved their scientific writing and deepened their critical thought, skills crucial not only to scientific careers but to an educated and functioning society.

Cian White

Natural Sciences Editor
Trinity Student Scientific Review 2017

TS
SR

Resurrection Plants:

The Evolution of Desiccation Tolerance and Potential Crop Improvement

Andrew Neill
Senior Sophister
Plant Sciences

Drought and water stress are severe limitations on crop productivity and under projected climate change these conditions are set to become more frequent across the globe, threatening up to 65% of agricultural land. Crop plants show no desiccation tolerance (DT) in their vegetative tissues despite having tolerant seeds. Resurrection plants are a diverse group of plants spanning the entire plant kingdom, except the gymnosperms, capable of surviving the loss of 95% of their relative water content (RWC). The genes responsible for this adaptation are based on the same gene families that code for DT in seeds. All seed-plants therefore possess these gene families within their genome, including crops, but they are only active in vegetative tissues in resurrection plants. These plants have evolved to utilise these genes within their vegetative tissues to overcome the oxidative and mechanical stresses of desiccation while maintaining macromolecular structures. Understanding this complex stress adaptation and its control within resurrection plants has potential to produce DT crops through genetic modification (GM). Recent research has focused on either incorporating resurrection plant genes into crop species or expressing seed DT genes within crop vegetative tissues. While this is a promising field of research and a potential solution to meeting world food demand, the lack of genetic data and field studies is limiting progress.

Agriculture Under Climate Change

Since the dawn of agriculture, crop species have been genetically modified through artificial selection for higher yields and disease and stress tolerance under prevailing climates. Modern human activity is changing the climate away from these optimal crop growing conditions.[1] Water stress and drought are predicted to become more frequent, last longer and set in faster than they currently do,[2, 3] and

the areas predicted to be most severely impacted by drought conditions include 80% of developing countries, which are unequipped to adapt quickly.[4]

Along with climate change, the world population is increasing. By 2050 the global population is predicted to reach 9 billion,[5] with most of this growth focused in developing countries. To meet the demands of this increased population, agricultural land must increase its output significantly. At least 65% of current agricultural land is threatened by drought,[4] and 1/3 of the world's population already live on land experiencing water stress, relying on irrigation.[6] Improving crops to withstand periods of water shortage is imperative for future food security and to meet increasing demand.

Nearly all angiosperms show desiccation (drying out) tolerance (DT), but it is restricted to seeds and pollen. This has been exploited by farmers for millennia; seeds such as Rice (*Oryza* spp.), Wheat (*Triticum* spp.) and barley (*Hordeum* spp.) can be dried and stored for long periods of time and remain viable. However, this trait is not present within vegetative tissues of traditional crops which are highly desiccation sensitive (DS). There is however, a group of highly diverse plants that can survive drying out to 5% of relative water content (RWC) (Fig. 1).[7] These plants are termed 'resurrection plants' because upon rewetting, they return to the vegetative state without damage. Understanding the mechanisms of this stress adaptation could result in a rich area for bioprospecting to improve traditional crops.

Figure 1: The hydrated (HD), Dry (water withheld for 14 days) and rehydrated (RD) state of the same *B. hygrometrica* plant. Scale bars represent 1 cm.[15]

Evolution of Resurrection Plants

The first macroalgae to successfully colonise land must have shown DT in order to survive terrestrial life,[7, 8] and DT is still common across the Bryophytes. DT was lost as tracheophytes evolved but appears again in the angiosperms. It has been proposed that the genes responsible for DT remained within the plant genome and were recruited into seed development within higher plants where DT seeds are

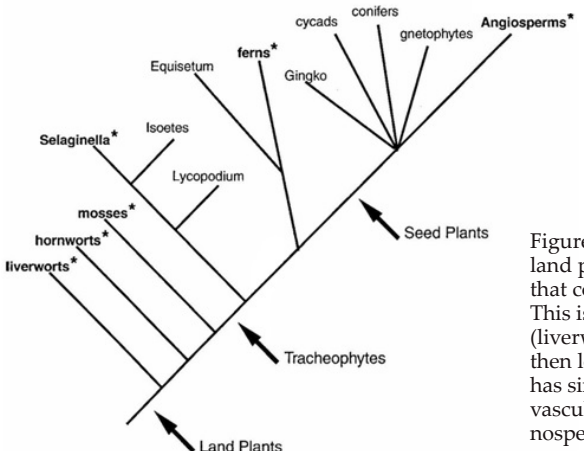


Figure 2: The evolutionary history of land plants with * denoting the groups that contain DT vegetative tissue (RPs). This is common to the Bryophytes (liverworts, hornworts, mosses) but was then lost as Tracheophytes evolved. It has since re-emerged within all major vascular plant groups, except the gymnosperms.[15]

almost a universal character.[9] Supporting evidence for this comes from the model plant *Arabidopsis thaliana* (L.) Heynh. in which 44% of seed DT genes have been shown to be "ancient" in their origin and well conserved across seed plants.[9]

Over 300 resurrection plants (RPs) have been identified within the higher plant groups. They are found within the ferns (Pteridophytes), spikemosses (*Selaginella sp.*) and angiosperms (Fig. 2) and the trait has evolved at least 12 times independently within the angiosperms (Fig. 3)[8-11]. RPs within the angiosperms alone have a distribution which spans the globe, including China, Australia, South America, Eastern Europe and are particularly numerous in South Africa suggesting this adaption is highly successful across a wide range of ecosystems and environmental conditions.[8, 11, 12]

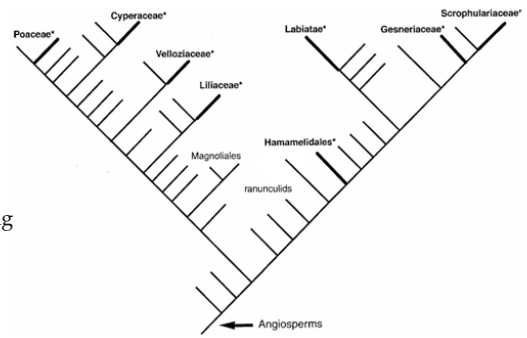


Figure 3: An angiosperm phylogeny showing lineages that contain RPs with *. These groups are unrelated indicating many independent origins of angiosperm vegetative DT.[10]

Angiosperm RPs show complex genetic history through high ploidy numbers and gene doubling events suspected of facilitating the repurposing of seed DT genes.[13, 14] Only one RP genome has been sequenced, the eudicot *Boea hygrometrica* (Bunge) R.Br., and no novel gene families were identified that could convey DT, consistent with the hypothesis that repurposed seed DT genes are responsible for vegetative DT character.[15] Therefore, RPs evolved DT from the repurposing of gene suites present within all seed-plants, but normally limited to seed tissue, and expressing them within vegetative tissue.[7, 9, 16]

Withstanding Desiccation

Why is water so essential for plant life? To summarise briefly, plant-water relations are complex but for vascular plants, are based around a continuous, unbroken column of water from the roots to the leaves. Water moves up this column along a water potential gradient composed of two major components: the osmotic potential (osmosis) and the pressure potential. Maintaining this gradient is essential to maintaining plant structure and support and to provide the water used in cellular pathways. A loss of water lowers both components, disrupting this gradient and the column of water can become broken by cavitation and wilting occurs rapidly.

A reduction to 5% RWC reduces the turgor pressure within cells and many mechanisms working synergistically are required to prevent death. There are two broad strategies employed by RPs to combat these stresses. The basal and more ancient Bryophytes invest in highly efficient, constitutive repair mechanisms upon rewetting to recover from damage while more advanced RPs show mechanisms induced by water stress that protect against damage rather than repair after damage.[12]

The stresses experienced by plant cells undergoing desiccation can be summarised into three major groups.[8, 12] Oxidative Stress (photooxidation) by reactive oxygen species (ROS) e.g. H_2O_2, $O_2^{\cdot-}$ produced inside the cell.[17] Mechanical stress caused by the reduction to 5% of RWC dramatically reduces cellular volume and turgor pressure that leads to rapid wilting, loss of structure and support and eventually the fusion or rupture of cell membranes. Finally, damage to macromolecules e.g. proteins, lipids as molecular structures are altered without aqueous media and leads to loss of function.

Oxidative Stress Adaptations

The major source of ROS within plant cells is the chloroplast. Under water stress, the stomata close, limiting CO_2 concentrations inside the leaf. The photosynthetic apparatus continues to receive light energy but instead of CO_2 fixation, ROS are produced. All RPs studied stop photosynthesising by 60% RWC.[16] Subsequent

response can be split into two groups:

Poikilochlorophyllous species, all monocot and some other angiosperm RPs, dismantle the thylakoid membrane upon drying and resynthesise upon rewetting. This completely prevents chloroplast ROS formation and these species can survive for over 10 months in the desiccated state.[8, 18] Homoiochlorophyllous species, most eudicot RPs, retain PSII and PSI during desiccation but deactivate their photosynthesising ability. This is initiated by the breakdown of the cytochrome b_6f complex preventing electron flow through the electron transport chain. By 60% RWC, the major photosynthesis protein complexes are arranged into a 2D lattice and photochemically inactivated. Antenna proteins e.g. LHCII remain present and can lower ROS concentrations.[19, 20] Upon rewetting, the photosystems are restored to their normal structure and reactivated.

Both strategies are highly effective at reducing ROS formation but homoiochlorophyllous species don't survive as long while desiccated, suggesting that they still produce some ROS.[8] Upon rewetting, homoiochlorophyllous species recover full photosynthetic capabilities within hours while poikilochlorophyllous species can take days.[21] The synthesis of new thylakoid membrane proteins causes this delay.[22] All RPs show a spike in respiration during rehydration for the repair and activation of the photosynthetic machinery.[22-24]

All plants are at risk of photooxidative damage and so antioxidants that quench ROS are universal across the plant kingdom.[17, 25-27] Many RPs upregulate these antioxidants due to the increased risk of ROS formation when undergoing desiccation. These molecules include anthocyanins, carotenoids, glutathione, catalases and superoxide dismutase (SOD) enzymes. Antioxidant concentrations tend to be higher in homoiochlorophyllous species, supporting the hypothesis that they are at an increased risk of ROS formation.[12] DT plants invest heavily into these antioxidants, e.g. 9% of the dry weight of *Haberlea rhodopensis* Friv. is myconoside, an antioxidant species, and this doubles upon drying.[23] One gene family, ELIP, shows significant extension in the genome of *B. hygrometrica* compared to DS relatives. ELIP is responsible for photoprotection and deactivating chlorophyll. Therefore, this RP appears to have evolved greater protection against photooxidation.[15, 28]

Mechanical Stress Adaptations

On the macro-scale, desiccating leaves curl inwards and leaf angle changes to shade inner tissues from the high temperatures and irradiances that drive ROS production.[29] This exposes only outer leaves to these stresses and RPs exposed to repeated cycles of wetting and drying often lose their outer leaves due to ROS damage while the shaded, inner leaves are protected and survive.[30]

Leaf and stem curling can also compensate for the reduction in cell volume and reduce mechanical stress. Another response to decreasing cell volume is a change

in cell membrane composition. A reduction in the saturation of phospholipids increases fluidity, allowing greater flexibility of the cell membrane under the mechanical pressure of cell shrinkage.[23] Other cells will begin to fold their cell walls to reduce cell volume and prevent plasmolysis (the complete separation of the cell membrane from the cell wall).[31] Cell wall fluidity is also increased by an increase in arabinose-rich polymers that act as carbon stores and improve cell plasticity.[32] In conjunction with increased fluidity, the cell membrane is often strengthened to resist rupture and tearing by the production of isoprene, a small and highly lipophilic molecule found across the plant kingdom.[33]

Vacuolisation is a conserved trait of RPs. This is the production of many small vacuoles that take up the majority of the cytoplasmic space under water stress.[28] The contents of these vacuoles is non-aqueous and speculated to include proline, sucrose, osmotic regulators (see below) and polyphenolic compounds.[34] This maintains cell integrity by providing a resistance to cell wall collapse in conjunction with cell folding.[12] The polyphenolics are thought to be photoreflectors, reducing light absorbance and ROS formation.[34]

Macromolecule Structural Protection

Without aqueous media, the highly specific protein structures within DS species become disrupted and aggregate together, destroying their function.[11] Molecules that help maintain molecular structure and function by regulating the osmotic potential of plant cells are called osmoregulators. All DT tissues, including seeds, show increased concentrations of sucrose, a non-reducing sugar and osmoregulator.[30, 35-37] Other sugars also increase inside RP cells but these are species specific responses e.g. octulose in *Craterostigma plantagineum* Hochst., trehalose in *Myrothamnus flabellifolia* Welw. and raffinose in *Xerophyta viscosa* Baker. These sugars can replace the hydrogen bonding of water *via* their hydroxyl groups (-OH) and stabilise macromolecules (Fig. 4).[38] They can also form inert intracellular glasses (vitrifiation) that maintain the structural integrity of cellular contents.[11, 24]

Raffinose is a particularly good osmoregulator due to its many hydroxyl groups and optimal glass formation has been shown to occur at a 5:1 sucrose-raffinose ratio.[36] Other specific sugar molecules produced include trehalose which is normally involved in carbohydrate signalling. *Tipogon loliiformis* (F.Muell.) Hubb., an Australian grass RP, has evolved a more complex role for this molecule. Under water stress, high concentrations of trehalose accumulate and trigger autophagy, a method of recycling cell contents e.g. salvaging N and C, and removing ROS.[39] While in most RPs this molecule acts as an osmoregulant, it has been repurposed in *T. loliiformis* in this complex stress signalling pathway.[13]

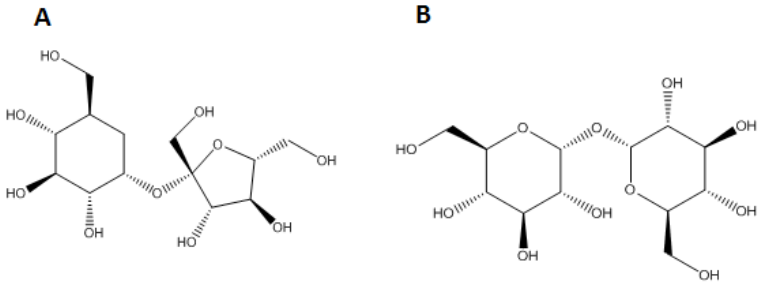


Figure 4: Chemical structure of A) – sucrose, B) – trehalose. The numerous hydroxyl groups facilitate hydrogen bonding. These sugars also allow glass formation within the cell, stabilising cell contents.

Can Resurrection Plants be Utilised in GM Crop Technology?

All crops are genetically modified from their ancestral species through artificial selection and already possess highly drought tolerant seeds e.g. Barley, Wheat, Corn (Zea mays L.), Rice.[7] When bioengineering crops for vegetative DT, it is vital that the aims are clearly defined. Is the goal to produce crops that can resist mild drought damage and continue being productive or crops that survive intense drought at the cost of reduced productivity?[40]

There are two approaches to utilising RPs for crop improvement:

Transforming crop plants using genes taken from RPs. This has been carried out in several studies. Overexpressing particular LEA (late embryogenesis) genes taken from RPs can lead to increased stress tolerance,[13] but this often leads to reduced productivity and growth rate.[41] Transgenic Tobacco, *Nicotiana tabacum* L., plants with RP LEA genes $BhLEA_1$ and $BhLEA_2$ retained 9% greater RWC under conditions of water shortage and showed less PSII damage.[31, 42] Placing an LEA gene from the RP *Cleistogenes songorica* (Roshev.) Ohwi, into *A. thaliana* caused accumulation of proline and improved viability under salt and drought stress.[43] Also from *C. songorica*, the gene $CsALDH_{12}A_1$ that codes for a detoxifying enzyme, improves survivability under stress of *A. thaliana*.[44]

Using mechanisms observed in RPs to alter the control of seed DT genes in crop vegetative tissues. Wheat (*Triticum aestivum* L.) seeds are able to reinduce DT in seeds 3 days after germination.[35] If this character could be extended to other tissues there is potential to produce DT wheat. Transcription factors taken from wheat seed tissue and placed into Tobacco cause upregulation of antioxidants and decreased photooxidative damage and therefore the regulation mechanisms of seed desiccation are well conserved.[45] The transciption factor family ABI_3 is involved in DT in seeds and RPs. Therefore the expression of this transcription factor family

has been altered in RPs to allow expression within vegetative tissue.[9] Mimicking this change through bioengineering could create drought tolerant crops.

Problems Faced in Producing Drought Tolerant Crops

The study of DT and its GM applications have seen an exponential increase since 1990,[46] and over 100 RP genes have been identified.[16] However, simply modifying or introducing one DT gene is not sufficient.[40] RPs have recruited whole gene families and rewired complex signalling networks to control DT in vegetative tissue and recreating this through GM technology will not be straightforward. Field studies are uncommon but laboratory conditions are not representative of how transgenic lines will perform under environmental conditions where water stress is unlikely to occur in isolation. Heat stress, salinity, nutrient availability and herbivore load must be taken into consideration.[40]

This research, while growing in interest, is also held back by the lack of genomic information of RPs. With only one species having its genome sequenced,[15] and wild populations showing inherent variation, pinning down genes and their regulatory mechanisms will prove challenging despite recent progress.[13] Often these DT genes show tight linkage to other traits that make predicting their effects in crop species difficult e.g. sucrose and trehalose are also used as signalling molecules in carbohydrate metabolism.[38]

The Future

The 'blue revolution' of agriculture has begun, aiming to reduce the amount of water required by crops to meet the yields required globally.[47] Current predictions show climate change occurring at a faster rate than plant species can naturally adapt and with increased drought, production is going to decrease unless action is taken. With the discovery of DT gene suites being present within all crop tissue, a new spark of research has begun, trying to use these genes, along with RPs, to create crops capable of withstanding drought. These RPs provide a rich area for bioprospecting and gene mining that could eventually lead to crop improvement. However, there is still a lot we do not understand about the regulation and control of RP DT genes and it is vital this becomes a central theme of research. Progress made within the past decade that creates an optimistic outlook into the field of DT crops.

References

1. IPCC. Climate change 2007- The Physical Science Basis: Working group I contribution to the Fourth Assessment Report of the IPCC: *Cambridge University Press*; 2007.

2. Dai A. Increasing drought under global warming in observations and models. *Nature Clim Change*. 2013;3(1):52-8.

3. Trenberth KE, Dai A, van der Schrier G, Jones PD, Barichivich J, Briffa KR, et al. Global warming and changes in drought. *Nature Clim Change*. 2014;4(1):17-22.

4. Zhengbin Z, Ping X, Hongbo S, Mengjun L, Zhenyan F, Liye C. Advances and prospects: biotechnologically improving crop water use efficiency. *Critical Reviews in Biotechnology.* 2011;31(3):281-93.

5. Cohen JE. Human population: the Next Half Century. *Science*. 2003;302(5648):1172-5.

6. Jogaiah S, Govind SR, Tran L-SP. Systems biology-based approaches toward understanding drought tolerance in food crops. *Critical Reviews in Biotechnology.* 2013;33(1):23-39.

7. Farrant JM, Moore JP. Programming Desiccation-tolerance: from Plants to Seeds to Resurrection Plants. *Current Opinion in Plant Biology*. 2011;14(3):340-5.

8. Gaff DF, Oliver M. The Evolution of Desiccation Tolerance in Angiosperm Plants: a Rare Yet Common Phenomenon. *Functional Plant Biology.* 2013;40(4):315-28.

9. Costa MCD, Farrant JM, Oliver MJ, Ligterink W, Buitink J, Hilhorst HMW. Key Genes Involved in Desiccation Tolerance and Dormancy Across Life Forms. *Plant Science*. 2016;251:162-8.

10. Oliver MJ, Tuba Z, Mishler BD. The Evolution of Vegetative Desiccation Tolerance in Land Plants. *Plant Ecology*. 2000;151(1):85-100.

11. Alpert P. Constraints of Tolerance: Why are Desiccation-tolerant Organisms so Small or Rare? *Journal of Experimental Biology*. 2006;209(9):1575.

12. Farrant JM. A Comparison of Mechanisms of Desiccation Tolerance among Three Angiosperm Resurrection Plant Species. *Plant Ecology.* 2000;151(1):29-39.

13. Leprince O, Buitink J. Desiccation tolerance: From Genomics to the Field. *Plant Science.* 2010;179(6):554-64.

14. Challabathula D, Bartels D. Desiccation Tolerance in Resurrection Plants: New Insights from Transcriptome, Proteome and Metabolome Analysis. *Frontiers in Plant Science.* 2013;4:482.

15. Xiao L, Yang G, Zhang L, Yang X, Zhao S, Ji Z, *et al*. The Resurrection Genome of Boea Hygrometrica: A Blueprint for Survival of Dehydration. *Proceedings of the National Academy of Sciences*. 2015;112(18):5833-7.

16. Moore JP, Farrant JM. A Systems-Based Molecular Biology Analysis of Resurrection Plants for Crop and Forage Improvement in Arid Environments. Improving Crop Resistance to Abiotic Stress: *Wiley-VCH Verlag GmbH & Co. KGaA;* 2012. p. 399-418.

17. Apel K, Hirt H. Reactive Oxygen Species: Metabolism, Oxidative Stress, and Signal Transduction. *Annual Review of Plant Biology*. 2004;55(1):373-99.

18. Bewley JD. Polyribosomes Conserved During Desiccation of the Moss Tortula Ruralis are Active. *Plant Physiology.* 1973;51(2):285-8.

19. Ingle RA, Schmidt UG, Farrant JM, Thomson JA, Mundree SG. Proteomic Analysis of Leaf Proteins During Dehydration of the Resurrection Plant Xerophyta Viscosa. *Plant, Cell & Environment*. 2007;30(4):435-46.

20. Zia A, Walker BJ, Oung HMO, Charuvi D, Jahns P, Cousins AB, *et al*. Protection of the Photosynthetic Apparatus Against Dehydration Stress in the Resurrection Plant Craterostigma Pumilum. *The Plant Journal.* 2016;87(6):664-80.

21. Challabathula D, Puthur JT, Bartels D. Surviving Metabolic Arrest: Photosynthesis During Desiccation and Rehydration in Resurrection Plants. *Annals of the New York Academy of Sciences*. 2015;1365(1):89-99.

22. Tuba Z, Protor CF, Csintalan Z. Ecophysiological Responses of Homoiochlorophyllous and Poikilochlorophyllous Desiccation Tolerant Plants: a Comparison and an Ecological Perspective. *Plant Growth Regulation.* 1998;24(3):211-7.

23. Moyankova D, Mladenov P, Berkov S, Peshev D, Georgieva D, Djilianov D. Metabolic Profiling of the Resurrection Plant Haberlea Rhodopensis During Desiccation and Recovery. *Physiologia Plantarum.* 2014;152(4):675-87.

24. Zhang Q, Song X, Bartels D. Enzymes and Metabolites in Carbohydrate Metabolism of Desiccation Tolerant Plants. *Proteomes*. 2016;4(4).

25. Mittler R. Oxidative Stress, Antioxidants and Stress Tolerance. Trends in Plant *Science.* 2002;7(9):405-10.

26. Mullineaux P, Karpinski S. Signal Transduction in Response to Excess Light: Getting out of the Chloroplast. *Current Opinion in Plant Biology.* 2002;5(1):43-8.

27. Triantaphylidès C, Krischke M, Hoeberichts FA, Ksas B, Gresser G, Havaux M, *et al.* Singlet Oxygen Is the Major Reactive Oxygen Species Involved in Photooxidative Damage to Plants. *Plant Physiology.* 2008;148(2):960-8.

28. Gechev TS, Dinakar C, Benina M, Toneva V, Bartels D. Molecular Mechanisms of Desiccation Tolerance in Resurrection Plants. *Cellular and Molecular Life Sciences.* 2012;69(19):3175-86.

29. Farrant JM, Vander Willigen C, Loffell DA, Bartsch S, Whittaker A. An Investigation into the Role of Light during Desiccation of Three Angiosperm Resurrection Plants. *Plant, Cell & Environment.* 2003;26(8):1275-86.

30. Yobi A, Wone BWM, Xu W, Alexander DC, Guo L, Ryals JA, *et al.* Comparative Metabolic Profiling between Desiccation-sensitive and Desiccation-tolerant Species of Selaginella Reveals Insights into the Resurrection trait. *The Plant Journal.* 2012;72(6):983-99.

31. Mitra J, Xu G, Wang B, Li M, Deng X. Understanding Desiccation Tolerance using the Resurrection Plant Boea Hygrometrica as a Model System. *Frontiers in Plant Science*. 2013;4:446.

32. Moore JP, Nguema-Ona EE, Vicré-Gibouin M, Sørensen I, Willats WGT, Driouich A, *et al.* Arabinose-rich Polymers as an Evolutionary Strategy to Plasticize Resurrection Plant Cell Walls against Desiccation. *Planta.* 2013;237(3):739-54.

33. Beckett M, Loreto F, Velikova V, Brunetti C, Di Ferdinando M, Tattini M, *et al.* Photosynthetic Limitations and Volatile and Non-volatile Isoprenoids in the Poikilochlorophyllous Resurrection Plant Xerophyta Humilis During Dehydration and Rehydration. *Plant, cell & environment.* 2012;35(12):2061-74.

34. Vander Willigen C, Pammenter NW, Mundree SG, Farrant JM. Mechanical Stabilization of Desiccated Vegetative Tissues of the Resurrection Grass Eragrostis Nindensis: Does a TIP 3;1 and/or Compartmentalization of Subcellular Components and Metabolites Play a Role? *Journal of Experimental Botany.* 2004;55(397):651-61.

35. Farrant JM, Bailly C, Leymarie J, Hamman B, Côme D, Corbineau F. Wheat Seedlings as a Model to Understand Desiccation Tolerance and Sensitivity. *Physiologia Plantarum.* 2004;120(4):563-74.

36. Peters S, Mundree SG, Thomson JA, Farrant JM, Keller F. Protection Mechanisms in the Resurrection Plant Xerophyta Viscosa (Baker): Both Sucrose and Raffinose Family Oligosaccharides (RFOs) Accumulate in Leaves in Response to Water Deficit. *Journal of Experimental Botany.* 2007;58(8):1947-56.

37. Yobi A, Wone BWM, Xu W, Alexander DC, Guo L, Ryals JA, et al. Metabolomic Profiling in Selaginella lepidophylla at Various Hydration States Provides New Insights into the Mechanistic Basis of Desiccation Tolerance. *Molecular Plant.* 2013;6(2):369-85.

38. Toldi O, Tuba Z, Scott P. Vegetative Desiccation Tolerance: Is It a Goldmine for Bioengineering Crops? *Plant Science.* 2009;176(2):187-99.

39. Williams B, Njaci I, Moghaddam L, Long H, Dickman MB, Zhang X, *et al.* Trehalose Accumulation Triggers Autophagy

During Plant Desiccation. *PLOS Genetics.* 2015;11(12):e1005705.

40. Ashraf M. Inducing Drought Tolerance in Plants: Recent Advances. *Biotechnology advances.* 2010;28(1):169-83.

41. Yang S, Vanderbeld B, Wan J, Huang Y. Narrowing Down the Targets: Towards Successful Genetic Engineering of Drought-Tolerant Crops. *Molecular Plant.* 2010;3(3):469-90.

42. Liu X, Wang Z, Wang L, Wu R, Phillips J, Deng X. LEA 4 Group Genes from the Resurrection Plant Boea Hygrometrica Confer Dehydration Tolerance in Transgenic Tobacco. *Plant Science.* 2009;176(1):90-8.

43. Zhang J, Kong L, Liu Z, Jahufer Z, Duan Z, Huo Y, *et al.* Stress-induced Expression Arabidopsis with a Dehydrin LEA Protein from Cleistogenes Songorica, a Xerophytic Desert Grass. *Plant Omics.* 2015;8(6):485.

44. Zhang J, Duan Z, Jahufer Z, An S, Wang Y. Stress-inducible Expression of a Cleistogenes Songorica ALDH Gene Enhanced Drought Tolerance in Transgenic Arabidopsis Thaliana. *Plant Omics.* 2014;7(6):438.

45. Hu WEI, Huang C, Deng X, Zhou S, Chen L, Li YIN, *et al.* TaASR1, a Transcription Factor Gene in Wheat, Confers Drought Stress Tolerance in Transgenic Tobacco. *Plant, Cell & Environment.* 2013;36(8):1449-64.

46. Lawlor DW. Genetic Engineering to Improve Plant Performance under Drought: Physiological Evaluation of Achievements, Limitations, and Possibilities. *Journal of experimental botany.* 2013;64(1):83-108.

47. Cominelli E, Tonelli C. Transgenic Crops Coping with Water Scarcity. *New Biotechnology.* 2010;27(5):473-7.

TS
SR

The Snowball Earth Hypothesis

Ellen Mullarkey
Senior Freshman
General Science

During Earth's history, evidence suggest that the Earth was completely frozen at least three times. These have been termed "Snowball Earth" events. Paleo-magnetic analysis from glacial deposits indicate deposition at low latitudes during the Neoproterozoic (1000-541 Million years ago (Ma)). Cap carbonates, negative carbon isotope anomalies and climate sensitivity also provide the most reliable evidence for the causes and terminations of a "Snowball Earth". Doubts have been raised on the significance of banded iron formations in proving atmospheric oxygen, with ongoing studies into the role of microbial organisms in their formation. Counter-evidence which favours a band of open water or a "Slushball Earth" includes climatic fluctuations which are impossible to achieve with a fully glaciated ocean and the ability for life to survive under such extreme conditions. This controversial hypothesis has prompted very interesting research and lines of evidence. Studying "Snowball Earth" events has aided current climate change research due to improved climate models.

Introduction

In 1948 Douglas Mawson discovered the glacial origin of sediment in the Elitina formation of south Australia and proposed a global Neoproterozoic glaciation.[1] Further evidence of low latitude glaciation prompted Brian Harland to suggest a global Neoproterozoic glaciation in 1964.[2] However, it was not until 1992 that paleomagnetist Joseph Kirschvink coined these periods as a "Snowball Earth" following studies pinpointing rock containing glacial evidence to low latitudes.[3] In order for this to be true he proposed three tests that needed to be proved; i) globally synchronous glacial sediment, ii) global climatic fluctuations and iii), ocean anoxia.[2,3] Hoffman *et al*. built upon the work of Kirschvink in the late 1990's,

popularising "Snowball Earth", particularly with evidence of cap carbonates which are believed to prove a greenhouse aftermath, linked to Kirschvink's second test.[4]

During the Neoproterozoic (~750 Ma) the fragmentation of the supercontinent Rodinia would have caused an increase in cyanobacteria due to an expansion of shallow continental shelves.[3] Due to the intensified photosynthesis, carbon dioxide would have been drawn down from the atmosphere and the planet subsequently cooled, creating a global glaciation.[3] The termination of "Snowball Earth" events would occur due to the build-up of carbon dioxide from volcanic activity and a subsequent greenhouse environment.[5] There is evidence for at least three global glaciations during the Neoproterozoic; The Makganyene (2.3 billion years ago (Ga)), the Sturtian (716.5 Ma), the Marinoan (635.5 Ma) .[6,7,9] They lasted at least 100 million years, 5 million years and 4 million years, respectively.[6,7,9]

Various analytical methods are used to support the "Snowball Earth" hypothesis. Low-latitude glacial deposits,[5,7] paleo magnetic data,[10,11] formation of cap carbonates,[4,5,12,13] negative carbon isotope anomalies,[5,14] occurrence of banded iron formations,[3,15] and climate sensitivity[4,16,17,18] all provide explanations for the possible causes and subsequent termination of a "Snowball Earth". Counter-evidence to the hypothesis is argued with evidence of climatic fluctuations during "Snowball Earths" and life's ability to survive such events.

Evidences of Low Latitude Glaciation

In order to confirm a global glaciation, it must be shown that the low latitudes were glaciated. Studies conducted in the Elitina formation in Australia, the Ghaub formation in Namibia and the Franklin large Igneous Province in north western Canada confirm the existence of low latitude ice-cover during the Marinoan and Sturtian glaciations.[5,7,10]

The Elitina formation displays glacial diamictite.10 Paleo magnetic studies of the Marinoan glaciation have shown that during the Elitina formation, Australia was located at ~8° latitude and underwent at least four polarity reversals. This indicates that other glacial deposits in Australia also lay below 30°.[10] The older Yaltipena formation is also shown to have occurred when Australia was at ~8.4°, and thus Australia was located at low latitudes prior to the Marinoan glaciation.[10] This is consistent with the rock record of an arid climate.[10] Polarity reversals indicate that the Marinoan may have lasted thousands to millions of years.[10]

The Congo craton in northern Namibia contains glacial deposits called the Otavi group, composed of the Ghaub and Chous formations. The Ghaub formation is composed of glacial diamictite, consisting of wackestone and clasts of carbonates. It is overlain by the Rasthof cap carbonate. The older Chous formation is composed of diamictite consisting of clasts of dropstones and carbonates. It is overlain by the Maieberg dolomite cap carbonate.[5] Pb-U zircon geochronology has shown the

Ghaub formation to be 635 ± 1.2 Ma.[8] Paleo magnetic studies by Meert and Van der Voo, indicate this group was at ~12°S at 735 ±30 Ma and ~39°S at 547 ±4 Ma.[19] This implies that the Otavi group was located at a relatively low paleolatitude during the Sturtian glaciation.

Evidence of Neoproterozoic glaciation in north western Canada is provided by inliers in Palaeozoic carbonates running from the Alaskan border, across Yukon and northwards to Victoria Island. A Sturtian aged inlier in the Ogilvie Mountains contains glacial diamictite composed of "bed-penetrating dropstones", "outsized clasts in fine, laminated beds and striated clasts"[7] that forms the Upper Mount Harper Group. Glacial push structures and soft sedimentary deformation indicate the presence of grounded ice. A volcanic tuff that erupted 716.5 Ma is interbedded with these glacial deposits and corresponds to the age of the Franklin Large Igneous Province. The Franklin Large Igneous Provence was part of north western Laurentia when the craton was within 10° of the equator,[11] confirming that grounded ice must have been present at this low latitude during the Sturtian glaciation.[7]

Carbon Isotope Ratios

Cap Carbonates, composed of dolomite (or locally limestone), overlie most Neoproterozoic glacial units and are not associated with regional scale glaciations.[4,5,12,13] Normally carbonates are precipitated in shallow tropical marine conditions and deposited in deeper water, with dolomite specifically containing magnesium usually from groundwater discharge. Their presence above glacial units signify the abrupt transition from glacial to tropical environments.[13] Thus dolomite and limestone provide evidence in support of a rapid change to greenhouse conditions and the following positive feedback albedo effect that ensued. As more water became available, it precipitated as acid rain while also causing intense weathering of glacial debris with run-off resulting in an increase in the alkalinity of seawater. Carbonate was then precipitated, filling the basins caused by the glaciation.[5] These cap carbonates all display low 13C values compared to other carbonates from the Neoproterozoic.[5]

Stable carbon isotope ratios of ^{12}C and ^{13}C can be used to provide evidence for the Sturtian and Marinoan glaciations. ^{12}C is an indicator of nutrient rich waters and indicates the biological productivity of an ocean. ^{12}C is sequestered into the sediment, leaving a positive $\delta^{13}C$ anomaly in the rock record. This means that a negative $\delta^{13}C$ anomaly in the rock record indicates a period of low biological productivity. In the Ghaub formation of the Otavi group of Namibia, $\delta^{13}C$ anomalies fell from 5-9 per mil to -5 per mil directly below the glacial surface, implying a decrease of biological productivity.[5] Directly after the glacial surface the anomaly dropped further to -6 per mil. This value is similar to the hydrothermal volcanic input of ^{13}C which supports the theory of a fully glaciated ocean with no atmosphere-ocean

gas exchange.[5] In the Windermere Supergroup of north-west Canada, the first 10m of limestones that cap the Rapitan tillites show a $\delta^{13}C$ decrease of 0 to -5 per mil.[14] Shields & Veizer have shown a strong global negative $\delta^{13}C$ anomaly of -22 per mil from ~600-800 Ma which correlates with the Sturtian and Marinoan glaciations, while a strong negative anomaly displayed at ~2000-2300 Ma broadly correlates to the Makganyene glaciation.[20] The negative anomalies provide evidence for three "Snowball Earth" events.

Planetary Oxidation

Diminished values of Mass Independent Fractionation (MIF) of Sulfur indicate the presence of free oxygen.[6] Conversely, in order to preserve MIF, there must be little free oxygen, sufficient sulfur gas and reducing gases.[21] Zahnle *et al*. show a sharp drop in the levels of MIF from 10-1 to 10-5 Tmol $year^{-1}$ at 2.45 Ga and propose that this was due to a decrease in methane rather than a rise in oxygen.[22] In contrast, Kopp *et al*. argue that a decrease of MIF at 2.09 Ga was due to the rise of free oxygen which destroyed a methane greenhouse.[6] They conclude that planetary oxidation commenced between the Horonian and Makganyene glaciations due to the rapid evolution of cyanobacteria. During glaciations there is a positive correlation between marine phosphate flux and continental chemical weathering which may have increased nutrient availability.[23] This could have destroyed the methane atmosphere resulting in freezing conditions.[6] In addition, the presence of free oxygen before the Makganyene glaciation is confirmed from red beds and hematitic paleosols with high sulfate content, derived from oxidation of sulfide.[6] Post snowball deposits of manganese such as The Kalahari Mn field suggest thriving cyanobacteria.[24] The precipitation of manganese in The Kalahari Mn field would require oxygen.[25] This implies that there was free atmospheric oxygen and a reduced amount of methane prior to the Makganyene glaciation. However more research is required to confirm the extent of the role which cyanobacteria had in lowering methane levels.

Banded Iron Formations

Banded Iron Formations (BIFs) have traditionally been used to provide evidence for free atmospheric oxygen. BIFs are thin-bedded or laminated chemical sediments, containing high amounts of iron.[15] Traditionally they were believed to have been deposited when ocean-atmosphere circulation of oxygen recommenced after a "Snowball Earth" leading to the oxidisation of ferrous iron which was released from mid ocean ridges.[3] However, many BIFs are contained within or below Neoproterozoic glacial deposits rather than lying above them, which has decreased confidence in the traditional view. Opponents to the traditional view have argued that BIFs may have formed from the oxidisation of ferrous iron to

ferric iron caused by microbial organisms[26] or by anoxygenic phototrophic bacteria prior to the evolution of photosynthesising cyanobacteria[27]. Although research is still ongoing it appears that BIFs are not as strong an indicator of atmospheric oxygen as was once thought.

Climate Sensitivity

Climate sensitivity is the way in which the global climate system responds to a given forcing.[28] Climate models simulate Earth's climate during the Neoproterozoic. A strong ice-albedo feedback is essential in order to support the snowball earth hypothesis. Sellers and Budyko were the first to use energy balance models to show this runaway ice-albedo feedback while further models have applied more precise constrains.[16] North *et al*. predict that if glaciation advances towards the equator past 30°- 40° latitude, then the ice-line will advance rapidly.[29] Sea ice has an albedo of 0.55 to 0.75 and so reflects 55% to 75% of incoming solar radiation, while snow covered sea ice reflects 65% to 80%.[30] An open ocean will only reflect ~6% of incoming solar radiation.[31] Voigt *et al*. show that a higher proportion of land in the tropics during the Marinoan glaciation compared to the present, would mean a more evenly distributed land surface area.[16] This would result in a higher surface albedo and consequently a cooling climate.[16] Furthermore, Pierrehumbert concludes that total solar irradiance was about 6% lower during the Marinoan than it is today[17] and thus it is likely that it was easier to facilitate low latitude glaciation. It is widely accepted that the Earth could have been launched into a full glaciation once a certain ice cover was reached.

Earth's climate sensitivity can also be viewed from the termination of snowball earth conditions. When a critical carbon dioxide value of 120,000 ppm is present in the atmosphere due to volcanic activity, the snowball earth environment would quickly turn into greenhouse conditions.[18] These greenhouse conditions would be further enhanced due to the positive feedbacks of an increasingly low albedo, as more land is exposed and more water vapour present in the atmosphere.[4] These studies of Earth's sensitivity to climate forcings support the "Snowball Earth" hypothesis by providing a plausible explanation for their termination.

Counter Evidence

The "Snowball Earth" hypothesis is still contentious with research ongoing. There has been wide speculation and arguments against a fully glaciated, hard, "Snowball Earth" and with many scientists favouring a "Slushball Earth" where some open water would exist. Rieu *et al*. suggests that there must have been some open ocean at the equator due to climate fluctuations.[32] It was also proposed that low latitude glaciation was caused by a reversed meridional climate gradient meaning that the equator would be cold, and the poles warm, if the earth's obliquity >54°.[33] However, Hoffman refuted this with the preferential precipitation of carbonates

in the tropics due to the carbonate compensation depth (CCD). Calcium is less soluble in warm water, which results in the CCD being deeper. Another criticism of a hard "Snowball Earth" is the argument that if life were to survive during a snowball event, and indeed life did survive, that earth could not have been fully glaciated. Moczydłowska states that an "open marine realm, active hydrologic cycle and free access to the seafloor"[34] were needed to maintain functioning organisms. McKay proposes that the survival of eukaryotic algae was due to their ability to photosynthesise because the ice may have been less than 10m thick with a transmissivity of >0.1%.[35] Others argue that many prokaryotic and eukaryotic organisms that were present in the Neoproterozoic are able to withstand severe environments and stress.[36,37] Consequently, the Neoproterozoic "Snowball Earth" would have acted as an environmental filter resulting in the Cambrian explosion,[14] which saw a rapid radiation in the number of species.

Extreme cyclic fluctuations has raised some doubts as to the extent of the "Snowball Earth" glaciations. During the late Neoproterozoic, Oman was located at ~13° latitude.[38] The Hufq supergroup of Oman contains the Fiq formation, consisting of both glacial and non-glacial marine sedimentary units. Rieu *et al*. have shown cyclic trends in the chemical index of alteration (CIA) of the Fiq formation which are due to climate change.[32] CIA values illustrate the extent of chemical weathering due to the removal of mobile cations. From the CIA record of the Fiq formation, glacial and interglacial events can be inferred. The data show three glacial periods with CIA values as low as 60 and interglacials with CIA values of up to 90. Rieu *et al*. argue that these extreme fluctuations cannot have been produced if the ocean was fully glaciated with a hydrological shutdown and a sublimation driven, reduced hydrological cycle.[32] If this is the case, it appears that there must have been some open water.

Application of Studying "Snowball Earth"

Studying snowball earth has enabled scientists to understand the interactions of ocean-atmospheric cycles and what they could mean for future climate change. Numerical climate models used to understand the causes and terminations of the "Snowball Earth" events, have stood as a benchmark for climate modelling. Models such as ECHAM5/MPI-OM have used sophisticated atmosphere-ocean general circulation models to study the initiation of "Snowball Earth" events as a result of albedo feedback.[16] By examining past albedo changes, perhaps results can be used to predict future feedback mechanisms concerned with carbon dioxide forcings in order to understand current climate change. Carbon dioxide concentrations rose from about 320 ppm to 403 ppm from 1950 to 2016.[39] Methane concentrations rose from about 12000 ppm to 17500 ppm from 1950 to 2012.[40] Using the results of "Snowball Earth" initiation and termination studies, it may be possible to predict if current rising greenhouse gas concentration trends could reach a tipping point with an albedo effect leading to the melting of all glacial ice.

Conclusions

There is strong evidence that there were three global glaciations or "Snowball Earth" periods in the Pre-Cambrian; the Makganyene, Sturtian and Marinoan glaciations. Although controversy exists about the presence of a band of open water, there is agreement that the rifting of Rodinia along with an increase the ratio of atmospheric oxygen to methane lead to freezing conditions and the onset of an albedo effect launching the earth into glaciation. The question of funding for non-practical research such as that of "Snowball Earth" is often posed, however it is clear that it is truly applicable to current climate studies.

Acknowledgements

This review was adapted from an essay "Snowball Earth" for GG2024 submitted in December 2016.

References

1. Mawson D. The late Precambrian ice-age and glacial record of the Bibliando dome. *J Proc R Soc New South Wales.* 1949;82:150–74.

2. Harland W. Critical Evidence for a Great Infra-Cambrian Glaciation. *Int J Earth Sci.* 1964;54(1).

3. J.L K. The Proterozoic Biosphere. Schopf J., Kleins C, editors. *New York: Cambridge University Press;* 1992. 51-52 p.

4. Hoffman PF, Kaufman AJ, Halverson GP. Comings and goings of global glaciations on a Neoproterozoic tropical platform in Namibia. *GSA Today.* 1998;8(5):1–9.

5. Hoffman PF, Kaufman AJ, Halverson GP, Schrag DP. A Neoproterozoic Snowball Earth. *Source Sci New Ser.* 1998;281(5381):1342–6.

6. Kopp RE, Kirschvink JL, Hilburn I a, Nash CZ. The Paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc Natl Acad Sci U S A.* 2005;102(32):11131–6.

7. Macdonald FA, Schmitz MD, Crowley JL, Roots CF, Jones DS, Maloof AC, *et al*. Calibrating the Cryogenian. *Science* (80-). 2010;327(5970):1241–3.

8. Hoffmann KH, Condon DJ, Bowring SA, Crowley JL. U-Pb zircon date from the Neoproterozoic Ghaub Formation Namibia: Constraints on Marinoan glaciation. *Geology.* 2004;32(9):817–20.

9. Prave AR, Condon DJ, Hoffmann KH, Tapster S, Fallick AE. Duration and nature of the end-Cryogenian (Marinoan) glaciation. *Geology.* 2016;44(8):631–4.

10. Sohl LE, Christie-Blick N, Kent D V. Paleomagnetic polarity reversals in Marinoan (ca. 600 Ma) glacial deposits of Australia: Implications for the duration of low-latitude glaciation in Neoproterozoic time. *Bull Geol Soc Am.* 1999;111(8):1120–39.

11. Park JK. Palaeomagnetic constraints on the position of Laurentia from middle Neoproterozoic to Early Cambrian times. *Precambrian Res.* 1994;69(1–4):95–112.

12. Kennedy MJ. Stratigraphy, Sedimentology, and Isotopic Geochemistry of Australian Neoproterozoic Postglacial Cap Dolostones: Deglaciation, 13C Excursions,

and Carbonate Precipitation. Vol. 66, *Journal of Sedimentary Research*. 1996. p. 1050–64.
13. Williams GE. Sedimentology, stable isotope geochemistry and palaeoenvironment of dolostones capping late Precambrian glacial sequences in Australia. *J Geol Soc* Aust. 2007;26(7–8):377–86.

14. Kaufman AJ, Knoll AH, Narbonne GM. Isotopes, ice ages, and terminal Proterozoic earth history. *Geology.* 1997;94:6600–5.

15. Klein C. Some Precambrian banded iron-formations (BIFs) from around the world: Their age, geologic setting, mineralogy, metamorphism, geochemistry, and origin. *American Mineralogist.* 2005;90(10):1473–99.

16. Voigt A, Abbot DS, Pierrehumbert RT, Marotzke J. Initiation of a Marinoan snowball earth in a state-of-the-art atmosphere-ocean general circulation model. *Clim Past.* 2011;7(1):249–63.

17. Pierrehumbert RT. High levels of atmospheric carbon dioxide necessary for the termination of global glaciation. *Nature.* 2004;429(6992):646–9.

18. Caldeira K, Kasting JF. Susceptibility of the early Earth to irreversible glaciation caused by carbon dioxide clouds. *Nature.* 1992;359:226–8.

19. Meert J, Van Der Voo R. Paleomagnetic and Study of the Sinyai Dolerite, Kenya: Implications for Gondwana Assembly. *J Geol.* 1996;104(2):131–42.

20. Shields G, Veizer J. The Precambrian marine carbonate isotope database: version 1.1. *Geochemistry Geophys Geosystems.* 2002;3(6):1–12.

21. Kump LR. The rise of atmospheric oxygen. *Nature*. 2008;451(7176):277–8.

22. Zahnle K, Claire M, Catling D. No Title. *Geobiology.* 2006;4(4).

23. Follmi KB. 160 my record of marine sedimentary phosphorus burial: coupling of climate and continental weathering under greenhouse and icehouse conditions. Vol. 23, *Geology.* 1995. p. 503–6.

24. Kirschvink JL, Gaidos EJ, Bertani LE, Beukes NJ, Gutzmer J, Maepa LN, *et al*. Paleoproterozoic snowball earth: extreme climatic and geochemical global change and its biological consequences. *Proc Natl Acad Sci U S A.* 2000;97(4):1400–5.

25. Liang M-C, Hartman H, Kopp RE, Kirschvink JL, Yung YL. Production of hydrogen peroxide in the atmosphere of a Snowball Earth and the origin of oxygenic photosynthesis. *Proc Natl Acad Sci U S A.* 2006;103(50):18896–9.

26. Konhauser KO, Hamade T, Raiswell R, Morris RC, Ferris FG, Southam G, *et al*. Could bacteria have formed the Precambrian banded iron formations? *Geology.* 2002;30(12):1079–82.

27. Widdel F, Schnell S, Heising S, Ehrenreich A, Assmus B, Schink B. Ferrous Iron Oxidation by Anoxygenic Phototrophic Bacteria. *Nature*. 1993;362(6423):834–6.

28. IPCC. IPCC Fourth Assessment Report (AR4). *IPCC* 2007;1:976.

29. North GR, Cahalan RF, Coakley JA. Energy balance climate models. Vol. 19, *Reviews of Geophysics.* 1981. p. 91–121.

30. Hibler III WD. A Dynamic Thermodynamic Sea Ice Model. Vol. 9, *Journal of Physical Oceanography.* 1979. p. 815–46.

31. NSIDC. Thermodynamics: Albedo [Internet]. Available from: https://nsidc.org/cryosphere/seaice/processes/albedo.html

32. Rieu R, Allen PA, Plötze M, Pettke T. Climatic cycles during a Neoproterozoic "snowball" glacial epoch. *Geology.* 2007;35(4):299–302.

33. Williams GE. Late Precambrian glacial climate and the Earth's obliquity. *Geol Mag.* 1975;112(5):441.

34. Moczydlowska M. The Ediacaran microbiota and the survival of Snowball Earth conditions. Vol. 167, *Precambrian Research.* 2008. p. 1–15.

35. McKay CP. Thickness of tropical ice and photosynthesis on a snowball Earth. *Geophys Res Lett.* 2000;27(14):2153–6.

36. Knoll a H. The early evolution of eukaryotes: a geological perspective. *Science.* 1992;256(5057):622–7.

37. Nealson KH. The limits of life on Earth and searching for life on Mars. *J Geophys Res.* 1997;102(E10):23,675-86.

38. Kilner B, Mac Niocaill C, Brasier M. Low-latitude glaciation in the Neoproterozoic of Oman. *Geology.* 2005;33(5):413–6.

39. Dlugokencky E, Tans P. No Title [Internet]. *Global CO2.* 2017. Available from: NOAA/ESRL(www.wsrl.noaa.gov/gmd/ccgg/trends/)

40. Pachauri R., Meyer L. IPCC, 2014: Climate Change 2014:Synthesis Report. *Contribution of Working Groups I,II,III to the fifth Assesment Report of the Intergovernmental Panel on Climate Change.* Geneva; 2014.

TS
SR

The Genomic Phenomenon: Polyploidy and its Contributions to Crop Improvement

Melike Arslan
Senior Sophister
Genetics

In recent years, our understanding of the structure, function and evolution of plant genomes has taken an explosive leap forward. Through DNA and whole genome sequencing, genome mapping and comparative genomic analysis, plant biologists have uncovered the molecular complexities that plant genomes contain. Throughout this review, the phenomenon of polyploidy (whole genome duplication) in plants will be discussed. The contributions made by polyploidy to plant genome evolution and consequential modelling of the modern plant genome structure will also be explored. Domestication of crop plants will be examined and how we can push the progress of crop improvement further through deepening our understanding of the processes involved in structuring the plant genome and its evolution.

Introduction

It is now strongly considered that a substantial amount of the diversity present in plants is a result of the duplication and adaptive modification of pre-existing genes.[1] Through comparative plant genomic studies, a key realisation emerged which outlines how plant gene families are significantly conserved, while taking into account the evolutionary timescales that include the diversification of all non-flowering plants and angiosperms.[1,2] A gene family is a group of functionally similar genes, based on shared nucleotide sequences, arising from the duplication of a progenitor gene. This datum indicates that major clades of the plant species have not produced new gene families during their evolutionary history but rather have possessed a basic genetic toolkit from the beginning.[1] Despite this, variations in gene family size which are specific to individual lineages are frequent among species. This indicates that the lineage-specific phenotypic variation and diversity present in plants may not be due to an equally diverse set of new genes. This hypothesis has lead to the commonly accepted theory that the processes of gene duplication and whole genome duplication (WGD), also referred to as polyploidy, play major roles in plant diversity. Polyploidy creates the raw material which is fundamental for adaptive evolution.[3] This phenomenon also, in part, explains why

plant genomes are exceptionally large in proportion to their relative sizes.[4]

The process of gene duplication is one that has largely captured the attention of plant biologists as no other taxa has a higher incidence of whole genome duplication.[1] The way in which genetic information is duplicated by an array of different autonomous mechanisms. Some of these processes include segmental and tandem duplications, frequently arising during DNA replication and recombination, but in particular, whole genome duplication events.[1] There are two types of polyploid plants, autopolyploid and allopolyploid. An autopolyploid organism is one whose genome consists of more than one chromosome set deriving from a single species (Figure 1). In this case, the chromosome doubling occurs spontaneously. In nature, autopolyploids occur due to the amalgamation of unreduced gametes but can also be produced artificially. Breeders of domesticated crop and commercial plants began the induction of chromosome doubling *in vitro* in order to provide stronger, more adaptable autopolyploid varieties.[5] Alternatively, they can seed naturally occurring polyploids, such as the *Miscanthus* polyploid complex which is an important bioenergy crop[6] These modified species are capable of withstanding harsher and more unfavourable environments and climates.[5] An allopolyploid genome is different, in that, it contains chromosome sets from different species and is a result of hybridisation (Figure 2). Some are sterile such as the allotriploid, *M. x giganteus*.[7]

Through the genomic sequencing of crops, our understanding of polyploidy has deepened and shows promise for crop improvement. The breeding of polyploid crop species has gained much attention since their domestication. While naturally occurring genetic improvements can be gained through recombination and selection, the most efficient and effective way of improving our crop species depends on a greater understanding of the mechanisms involved, but also, uncovering the intricacies of genetic variation which underlie a phenotype.

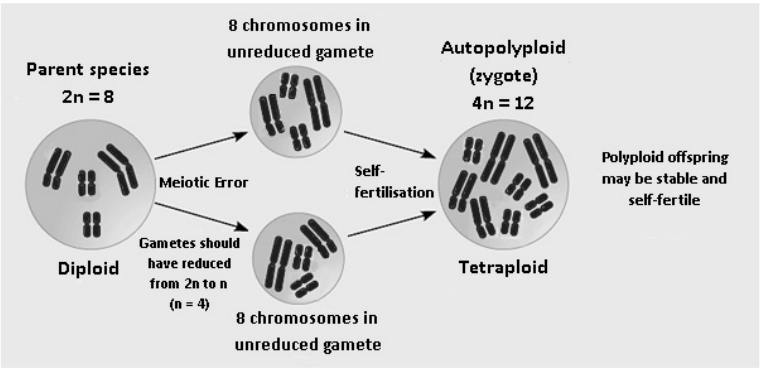


Figure 1: The induction of an autopolyploid. An autopolyploid is an organism whose chromosome number is doubled from 2n to 4n due to meiotic mutations, followed by self-fertilisation. In effect, it is a whole genome duplication.
Figure adapted from Chapter 24, 2008 Pearson Education Inc, publishing Benjamin Cummings.

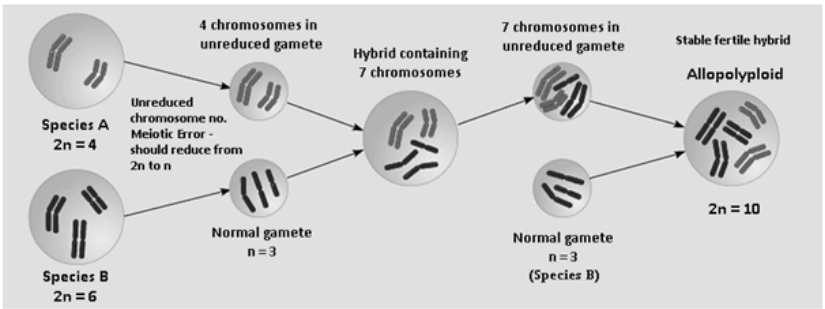


Figure 2: The induction of an allopolyploid. An allopolyploid is an organism with genome duplication whose chromosomes are derived from two separate species. The figure shows one mechanism by which this can happen. This polyploidy event involves an unstable hybrid intermediate, however, once it has been crossed with a normal gamete the result is a stable fertile allopolyploid hybrid.
Figure adapted from Chapter 24, 2008 Pearson Education Inc, publishing Benjamin Cummings.

Polyploidy: its Ubiquity and Cyclical Nature

Much of the genomic studies over the past fifteen years have shown that all angiosperms (flower and seed producing plant species) are polyploid. Furthermore, comparative genome mapping suggests that the angiosperm phylogenetic and molecular evolution contains multiple cycles of polyploidy.[8] That is, the most recent polyploidization over lays earlier duplications which occurred in the beginning of angiosperm evolution and previous to that, duplications took place at the end of the *Spermatophytina* (seed plants) period. The studies involving expressed sequence tags have uncovered 'peaks' of parallel sequences between genes within genomes, signifying a number of gene duplicates.[8] Not only that, their shared existence and characteristics imply a common ancestor.[9] In addition, a number of such peaks were present within individual genomes, suggesting more ancient polyploidy events. Studies done on *Arabidopsis thaliana* (L.) Henynh[10, 11] have confirmed that a number of polyploidy events occurred during this species' evolution, providing strong evidence that one cycle of genome duplication took place after the eudicot divergence and a second WGD event occurred at some stage after *A. thaliana* and *Brassica* genus diverged from their common ancestor. The results of these studies are consistent with the generally accepted hypothesis that *A. thaliana* possesses a particularly minimum angiosperm genome and yet even its small genome has been subjected to the cyclical nature of polyploidy.[12] Figure 3 demonstrates the wealth of species that also exhibit repeated genome duplication events during their evolution.

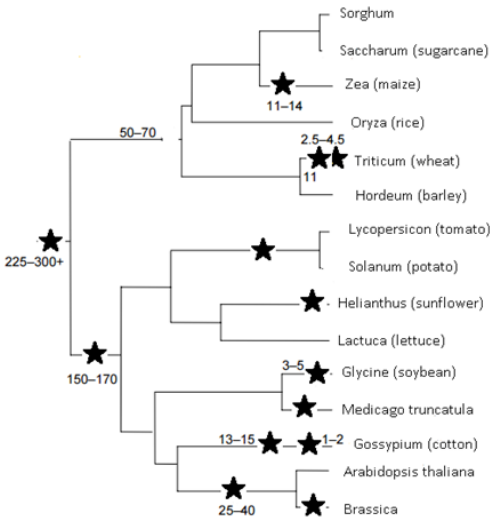


Figure 3: The evolution of angiosperms and the inferred polyploidy events that occurred throughout. The black stars represent presumed large-scale duplication events. The numbers in the figure are roughly estimated dates (in millions of years) since the WGD event. The branch lengths are not to scale.
Figure adaptation[13]

The Diversification Following Polyploidy Events

Something to consider when discussing gene duplication, but not necessarily with other forms of mutation, is that this process produces genetic redundancy. That is, it produces a number of genes within the one genome that perform the same function. It is accepted that following these polyploidy events, the duplicated genes experience a period of relaxation from selection and so the duplicates are given the opportunity to explore new evolutionary paths.[14] However, a large proportion but not all of the genetic redundancy is removed through the process of gene loss. An excellent example of how gene duplication is followed by the process of differential gene removal is Maize, *Zea mays* L. ssp. mays. Studies done on this crop plant species that behaves meiotically like a diploid (two sets of chromosomes) but is an ancient polyploid,[15-17] demonstrate that approximately half of all duplicate genes have been removed in the estimated 11 million years since the duplication event that produced the progenitor of Maize.[14] On an even grander scale of evolutionary time, an illustrative example would be that of the divergence of angiosperms. The twin processes of polyploidy and gene loss has resulted in the creation of modern angiosperm genomes. These genomes are exceptional in that they present hierarchical networks of syntenies with only partial gene membership of any one single linkage group.[14,17,18]

The genetic redundancy created following genome duplications has been thought to provide evolutionary opportunities due to the subsequent relaxation of purifying selection on the duplicated genes. Ohno classically formulated in his book 'Evolution by gene duplication' (1970), that when given sufficient time, one copy of a duplicated gene can obtain a beneficial mutation which may result in the preservation of both copies.[14] This theory is commonly referred to as neofunctionalisation. An alternative theory is subfunctionalisation which can arise in the absence of natural selection.[14] This concept explains that mutations may collect among gene duplicates in such a manner that they separate or 'partition' cumulative ancestral functions. Both sets of duplicates that result from this are maintained in order to complete their corresponding ancestral functions.

Over the past ten years, research into the consequences following polyploidy has been revealing. It seems as though; genome-wide subfunctionalisation and neofunctionalisation are representative of long term genomic responses,[19-21] as well as significant genomic structural rearrangements. Reduced chromosome numbers and the loss of repetitive sequences and duplicates, on a large-scale are consequences of these structural rearrangements.[8,22,23] This means that new polyploid species, having been subjected to multiple rounds of polyploidization, in time, experience enormous loss of redundant DNA and periodic genome downsizing.[24] This therefore explains how neopolyploid species (most recent polyploid) in the end become diploidized by these mechanistically varied processes. The modern descendants behave more and more as natural diploid species, all the while, concealing the vestigial evidence of multiple ancestral WGD events within their genomes.[8] In simpler terms, all contemporary angiosperm lineages are paleopolyploid (genome containing WGDs that occurred at least several million years ago), only they differ in the number of episodes of WGD events in their evolutionary history and in the timing of these events.

Polyploidy, Domestication and the Future Crop Improvement

A large proportion of crop species are recent allopolyploids. Therefore, an in depth knowledge of the evolutionary outcomes of gene duplication is not only fundamental to our understanding of angiosperm evolutionary history, it is also vital for agricultural domestication and improvement.

Many evolutionary processes affect the structure and function of polyploid genomes such as, transgressive gene expression alterations, genome reorganisation, genome downsizing and as mentioned earlier, subfunctionalisation and neofunctionalisation of duplicate genes.[5] In most cases, genetic alterations like these are accompanied by heterosis which is the improved function of a biological phenotype present in hybrid progeny.[25] Heterotic offspring usually exhibit robustness, vigour and an increase in crop yield compared with their diploid relatives. It isn't surprising then, that there has been a progressive interest in

heterotic polyploid crop plants since the initial domestication of crop plants.[26-29] A network of phenotypes in domesticated crops which are distinguishable from those of their wild ancestors include, a loss of dormancy, increases in biomass, growth rate and fertility, enhanced palatability, a shift from the perennial to annual life cycle and a wider geographical range.[5,30,31]

The diversity in wheat, *Triticum*, grains demonstrates the extent of phenotypes important in the domestication process, which has been facilitated by polyploidy.[5,30,32] Triticum species are part of a polyploid complex with diploids (e.g. *T. monococcum* L.; Einkorn), tetraploids (e.g. *T. durum* Desf.; pasta wheat) and hexaploids (*T. aestivum* L.; bread wheat). Grain hardness is controlled by the *Hardness* (*Ha*) locus in polyploid wheat species.[5,32] The locus' evolutionary history illustrates the contributions polyploid provide crop improvement. The locus present in diploid wheat species consists of a number of closely linked genes which collectively, confer the soft grain phenotype. While, in allotetraploid wheat, *T. durum*, these genes have been removed from the genome and consequently resulted in its hard grain phenotype which is ideal for pasta production.[32] To further the domestication process, allotetraploid wheat when crossed with a D-genome diploid wheat, results in an allohexaploid wheat, *T. aestivum*, with a soft grain phenotype, bread wheat.[5] The soft grain phenotype was restored through the introduction of the D-genome as it contains the *Ha* locus which is absent in the parent allotetraploid. In other words, under diversification by selection and at various times in different cultural contexts, allopolyploidization in wheat has facilitated the removal and reintroduction of the soft grain phenotype,[5,30] providing a large variety of wheat species, ultimately contributing to the domestication of crop plants.

Although significant advances have been made in the domestication process and crop improvement, a topic of discussion is where the future of crop improvement is heading. An ongoing challenge in crop improvement is the sequencing of the polyploid genomes.[33,34] In order to fully understand the inner genomic workings of a plant species, one must produce a 'plan' of the species' genomic structure which is achieved through various sequencing techniques.[33] However, compared to diploid species, polyploids impart more experimental challenges which, for the most part, are due to their genome multiplicity.[28] Firstly, in comparison to their diploid relatives, polyploid genomes are much larger and so, for this reason alone, makes them less desired as model organisms for study. Furthermore, it is far more costly to sequence larger genomes and the level of computational skill required for genomic assessment is far greater than that of the preferred inbred diploid relatives.[28]

Neopolyploid species have been subject of much crop improvement research as they have a rich domestication history and exhibit an abundance of accumulated phenotypes useful for crop improvement. However, they present particular challenges for this analysis, due to their high genic and repeat similarity between their constituent genomes.[28] With that in mind, it is anticipated that these current challenges will soon be addressed by advances in genome sequencing.[35]

In addition, it is very possible that epigenetic variation which has been induced by polyploidy and hybridisation contributes to heterosis. This was demonstrated in recent studies done on inbred *A. thaliana* plants, the data showing that epigenetic variability within a population may result in increased crop yield.[36] This epigenetic characteristic of heterosis is an exciting discovery as it will allow the use of epigenetic diversity with crop plants and ultimately provide a novel opportunity for crop improvement in the future.

Conclusions

Few would dispute that gene duplication is a major driver of plant functional diversity. If one was to consider this idea logically, it is far easier to create new phenotypes or functions from pre-existing genes rather than starting from scratch. Upon genome doubling, each copy has the opportunity to evolve independently whereby the outcome may lead to some functional novelty within the species. Today, due to the work done in countless studies it is more appropriate to ask 'when the WGD event occurred and how many cycles of duplication took place during the species' evolution?' rather than questioning whether or not the species is polyploid. With further research into the fundamental architecture of the plant genome, in the hopes of fine tuning our knowledge on the processes that guided the evolution of our modern crops, we should be able to utilize it in a way that is beneficial to humankind. That is, developing crops that can withstand harsher or unfavourable environments and climates, standardising crop productivity across the world. It is now accepted that since the beginning of the agricultural age, the domestication of wild plants was the fundamental innovation which ultimately led to the expansion of the human population.[37] Tracing back to cultures long lost, plant domestication and continued crop and agricultural improvement by means of artificial selection, produced the new phenotypes that, to this day, sustain human populations across the world.

Acknowledgements

This work was adapted from a review: "Plant genomes: Structure and function" for the course GE4013 (Genetics Literature Review). The author would like to thank Mr C. White and Dr T. Hodkinson for helpful appraisals.

References

[1]Flagel LE, Wendel JF. Gene duplication and evolutionary novelty in plants. *New Phytol*. 2009;183(3):557-564.

[2]Rensing SA, Lang D, Zimmer AD, *et al*. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science*. 2008;319(5859):64-69.

[3]Ni Z, Kim ED, Ha M, *et al*. Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature.* 2009;457(7227):327-331.

[4]Puttick MN, Clark J, Donoghue PC. Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proc Biol Sci.* 2015;282(1820):20152289.

[5]Renny-Byfield S, Wendel JF. Doubling down on genomes: polyploidy and crop plants. *Am J Bot*. 2014;101(10):1711-1725.

[6]Hodkinson TR, Klaas M, Jones MB, Prickett R and Barth S. Miscanthus: a case study for the utilization of natural genetic variation. *Plant Genet. Resour.* 2015; 13(3)219–237

[7]Hodkinson TR, Chase MW and Renoize SA. Genetic resources of Miscanthus. *Asp Appl Biol*. 2001; 65:239-258

[8]Wendel JF, Jackson SA, Meyers BC, Wing RA. Evolution of plant genome architecture. *Genome Biol.* 2016;17:37.

[9]Vanneste K, Van de Peer Y, Maere S. Inference of genome duplications from age distributions revisited. *Mol Biol Evo*l. 2013;30(1):177-190.

[10]Grant D, Cregan P, Shoemaker RC. Genome organization in dicots: genome duplication in Arabidopsis and synteny between soybean and Arabidopsis. *Proc Natl Acad Sci U S A.* 2000;97(8):4168-4173.

[11]Ku HM, Vision T, Liu J, Tanksley SD. Comparing sequenced segments of the tomato and Arabidopsis genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *Proc Natl Acad Sci U S A.* 2000;97(16):9121-9126.

[12]Blanc G, Wolfe KH. Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *Plant Cell.* 2004;16(7):1679-1691.

[13]Adams KL, Wendel JF. Polyploidy and genome evolution in plants. *Curr Opin Plant Biol.* 2005;8(2):135-141.

[14]Ohno S. 1970. Evolution by gene duplication. New York, NY, USA: Springer-Verlag

[15]Lai J, Ma J, Swigonová Z, *et al*. Gene loss and movement in the maize genome. *Genome Res*. 2004;14(10A):1924-1931.

[16]Messing J, Bharti AK, Karlowski WM, *et al*. Sequence composition and genome organization of maize. *Proc Natl Acad Sci U S A*. 2004;101(40):14349-14354.

[17]Ilic K, SanMiguel PJ, Bennetzen JL. A complex history of rearrangement in an orthologous region of the maize, sorghum, and rice genomes. *Proc Natl Acad Sci U S A.* 2003;100(21):12265-12270.

[18]Paterson AH, Bowers JE, Peterson DG, Estill JC, Chapman BA. Structure and evolution of cereal genomes. *Curr Opin Genet Dev.* 2003;13(6):644-650.

[19]Liu SL, Adams KL. Dramatic change in function and expression pattern of a gene duplicated by polyploidy created a paternal effect gene in the Brassicaceae. *Mol Biol Evol.* 2010;27(12):2817-2828.

[20]Liu Z, Xin M, Qin J, *et al*. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (Triticum aestivum L.). *BMC Plant Biol.* 2015;15:152.

[21]Hughes TE, Langdale JA, Kelly S. The impact of widespread regulatory neofunctionalization on homeolog gene evolution following whole-genome duplication in maize. *Genome Res.* 2014;24(8):1348-1355.

[22]Renny-Byfield S, Gallagher JP, Grover CE, *et al.* Ancient gene duplicates in Gossypium (cotton) exhibit near-complete expression divergence. *Genome Biol Evol.* 2014;6(3):559-571.

[23]Woodhouse MR, Cheng F, Pires JC, Lisch D, Freeling M, Wang X. Origin, inheritance, and gene regulatory consequences of genome dominance in polyploids. *Proc Natl Acad Sci U S A.* 2014;111(14):5283-5288.

[24]Leitch AR, Leitch IJ. Genomic plasticity and the diversity of polyploid plants. *Science.* 2008;320(5875):481-483.

[25]Schnable PS, Springer NM. Progress toward understanding heterosis in crop plants. *Annu Rev Plant Biol.* 2013;64:71-88.

[26]Hua J, Xing Y, Wu W, *et al*. Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid. *Proc Natl Acad Sci U S A.* 2003;100(5):2574-2579.

[27]Riedelsheimer C, Czedik-Eysenberg A, Grieder C, *et al*. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat Genet.* 2012;44(2):217-220.

[28]Comai L. The advantages and disadvantages of being polyploid. *Nat Rev Genet.* 2005;6(11):836-846.

[29]Mohammadi SA and Prasanna BM. Analysis of Genetic Diversity in Crop Plants— Salient Statistical Tools and Considerations. *Crop Sci.* 2003; 43:1235-1248. doi:10.2135/cropsci2003.1235

[30]Olsen KM, Wendel JF. A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol.* 2013;64:47-70.

[31]Salamini F, Ozkan H, Brandolini A, Schäfer-Pregl R, Martin W. Genetics and geography of wild cereal domestication in the near east. *Nat Rev Genet.* 2002;3(6):429-441.

[32]Chantret N, Salse J, Sabot F, *et al*. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (Triticum and Aegilops). *Plant Cell.* 2005;17(4):1033-1045.

[33]Olsen KM, Wendel JF. Crop plants as models for understanding plant adaptation and diversification. *Front Plant Sci.* 2013;4:290.

[34]Edwards D, Batley J, Snowdon RJ. Accessing complex crop genomes with next-generation sequencing. *Theor Appl Genet.* 2013;126(1):1-11.

[35]Faino L, Thomma BP. Get your high-quality low-cost genome sequence. *Trends Plant Sci.* 2014;19(5):288-291.

[36]Latzel V, Allan E, Bortolini Silveira A, Colot V, Fischer M, Bossdorf O. Epigenetic diversity increases the productivity and stability of plant populations. *Nat Commun.* 2013;4:2875.

[37]Harari, Yuval N. 2015. Sapiens: a brief history of Humankind. *New York, USA: Harper*

TS
SR

The Rhino Horn Trade:

Opportunities and Implications for the Legalised Sale of Bio-fabricated Rhinoceros Horn. Can the Current Poaching Crisis be Halted?

Eleanor Mullen
Senior Freshman
General Science

Developments in synthetic biology have made the bio-fabrication of rhinoceros horn possible. Pembient, an IndieBio Seattle based company, purveyor of bio-fabricated horn, intends to sell bio-fabricated substitution products at a fraction of the price of wild sourced horn, thereby making poaching unprofitable. A number of conservationist and governmental bodies are concerned that this will instead increase demand for rhino horn, making the illicit market harder to control. Petitions such as those made by Wild Aid and the World Centre of Biodiversity have attempted to ban the introduction of bio-fabricated horn to the market. This review examines the conservational impact of legalizing bio fabricated horn and possible policy reviews for rhino conservation.

Introduction

The rhino (Family *Rhinocerotidae*), a mammal that once roamed over much of Eurasia and South Africa, now struggles to survive in national reserves. There are five species of rhino. The International Union for the Conservation of Nature (IUCN) Red List of threatened species categorises two of the Asian rhinos, the Sumatran Rhino (*Dicerorhinus sumatrensis* G. Fischer, 1814) and Javen Rhino (*Rhinoceros sondaicus* Desmarest, 1822) as critically endangered.[1] With limited gene pools and relentless poaching, both populations are in rapid decline.[2] The third Asian rhino, the Greater One-horned Rhino (*Rhinoceros unicornis*, Linnaeus, 1758) is listed as vulnerable on the IUCN Red List.[1] African rhino species are also under imminent threat. Africa's Black Rhino (*Diceros bicornis* Linnaeus, 1758) is categorised as "Critically Endangered",[1] only 4,880 remain in the wild.[3] Africa's larger White Rhino (*Ceratotherium simus* Burchell, 1817) numbers 20,165 in the wild and is listed "Near Threatened".[1,3]

Between the 1960s and the 1990s rhino populations plummeted as a result of relentless poaching. In the 1990s, international action caused poaching levels to effectively halt.[4] Rhino populations began to recover until 2008, when a newly emerging market between Vietnam and Africa lead to the current poaching crisis.[4] South Africa, the main trader with Vietnam, has borne the brunt of this crisis with the number of both Black and White Rhinos poached per year increasing from 13 in 2007 to 83 in 2008.[4] The poaching crisis is a result of the increased demand for rhino horn. The retail price of rhino horn has seen a tremendous increase: from $4,700 per kg in 1993 to $65,000 per kg in 2012.[5] Primary end use markets include Vietnam, China and Southeast Asia.[6] The increased demand for rhino horn appears to correlate with economic development in Asia, and associated elevated levels of disposable income.[7] Motives for rhino horn purchases fall into three main categories: investment and collectible value, artistic value as a fine art carving material that can be crafted into functional and ornamental items,[3] and medicinal use such as reducing internal heat and removing toxins from the blood.[5]

Conservation initiatives have failed to halt the escalation of the current poaching crisis and have proved insufficient in preventing the extinction of the Western Black Rhino (*Diceros bicornis longipes* Zukowsky, 1949).[4] Poachers are becoming ever more resourceful and are making use of online information, efficient networking and better weapons.[8] Poaching is evolving, and so conservation must evolve faster.[8] Current strategies must be examined and reviewed.

Bio-fabricated Horn

One new strategy for rhino conservation is supply enlargement through the manufacturing of substitution products. Synthetic biology, the engineering of biological based systems, can be used to produce bio-fabricated rhino horn in the laboratory.[9] Rhino horn consists mainly of keratin, lacking the bony horn core that is commonly present in the horn structure of most ungulates.[10] One method of rhino horn bio-fabrication is to engineer yeast cells to produce the same keratins found in wild rhino horn.[9] These keratins are then combined with other natural components of rhino horn, including trace elements and rhino DNA.[9] A 3D printing process, involving successive layering of bio fabricated product, can be used to synthesise the three-dimensional structure of rhino horn.[11] The technique requires development and refinement in order to achieve the physical, genetic, and spectrographic properties present in wild horn.[9] One company developing this technology is the Seattle based company, Pembient.[12]

Discussion and Analysis of the Introduction of Bio-fabricated Horn to the Market

Pembient intends to target the carving industry, marketing their products as certified bio-fabricated horn. The company expects that after bio-fabricated horn is resold, product trends will follow Akerlof's "Market for Lemons".[12] The underlying principle of The Market for Lemons is that asymmetric information effects the value of goods on the market.[13] An asymmetric market is one in which the customer and seller have an unequal amount of information about the commodity being sold.[13] Pembient predicts that rhino horn buyers will consider bio-fabricated horn less valuable than wild sourced rhino horn. The customer won't know whether they are purchasing authentic, valuable rhino horn or less valuable bio-fabricated horn. Customers will be unwilling to pay high prices when there is a risk that the horn they wish to purchase is bio-fabricated. The price of buying wild sourced rhino horn then decreases. There is still a profit in selling bio-fabricated horn, but less profit in poaching wild sourced horn. The cost of producing synthetic products decreases as technology improves. The price of wild horn further decreases, poaching becomes unprofitable and not worth the risk.[12] The application of Akerlof's model to the bio-fabricated horn market is, however, currently unsupported by empirical data.[7] There is concern that the sale of bio-fabricated horn will not follow Akerlof's model, and will instead increase desire for wild sourced rhino horn.[14]

Marketing Concerns

Conservation efforts have invested considerable resources into demand reduction efforts, spreading scientific knowledge to inform the end-user that rhino horn has no medicinal properties.[9] Allowing the sale of bio-fabricated horn means promoting supply enlargement. Rhino horn buyers may view the sale as an admission that rhino horn does have medicinal properties.[9] Perhaps bio-fabricated horn could be marketed as an ethical substitution product that has no medicinal value. This kind of marketing has worked in the fur industry - an example being Zulu churches in South Africa endorsing the use of synthetic leopard pelts.[15] A substitute product led to supply enlargement, and an alternative to leopard poaching.[15] There is also concern that the economic incentives of poachers will not be addressed.[14] However, the production of substitution products could present alternative, less dangerous employment for poachers, thus reducing economic incentive for poaching.[16]

Another concern is that the sale of bio-fabricated horn may severely complicate law enforcement and facilitate reverse laundering.[9] Individuals in possession of poached horn may claim that they thought their horn was bio-fabricated. This could be used as a defence against law enforcement. Pembient intend to use cryptography and DNA water marking technologies to identify their product and combat reverse laundering.[9] This supports the work of the South African Government, in partnership with CITES and UNEP, who are working to improve

Rhino horn DNA testing and forensics capacity, along with Veterinary Genetic Laboratories and protocols for wildlife crime science investigation.[17] This aids conservation genetics and gene banking, also boosting law enforcement capacity, making crime investigations more effective.[18] Laws governing rhino horn trade will require review if the sale of bio- fabricated horn is legalised, as this is expensive and challenging to enforce. However, reviewing these laws may simultaneously address high levels of corruption in law enforcement and cross border communication in apprehend criminals.[6,18]

Consumer demand

Little is known about how bio-fabricated horn will be viewed by buyers of illegal horn.[7] A small number of surveys have been conducted, and various predictions have been made.[7] One such survey was a choice experiment, undertaken in Vietnam with a sample size of 850 residents, to assess how demand for rhino horn varies depending on its source.[5] Horn from wild rhinos is believed to have greater medicinal properties than that of captive or farmed rhinos.[5] This poses many questions about how bio-fabricated horn will be viewed. If the DNA used in the production of bio-fabricated horn is taken from the most desirable wild rhino, would it be as desirable as wild sourced horn? Would it be more desirable than farmed rhino horn? There are already counterfeit rhino horns on the market, carved from alternative materials such as Water Buffalo (Family *Bovidae*) horn. Will bio-fabricated horn be viewed as another counterfeit despite being indistinguishable from wild sourced horn? In addition, it has been found that the majority of rhino horn consumers prefer horn acquired from non-lethal methods of harvesting.[5] Could this make bio-fabricated horn more desirable? Purchasing from the illicit market has risks, for instance the horn could be sourced from a museum in Europe, or could be poisoned. Will buyers of rhino horn for medicinal use view bio-fabricated horn as a safer product? In order to determine whether or not bio-fabricated horn will be beneficial to rhino conservation, these questions need to be addressed. Current knowledge of the rhino horn market is insufficient.[7]

Pembient intends to market rhino horn at a fraction of the price of wild-sourced rhino horn.[12] This might extend the market to which rhino horn is available, encompassing a larger economic bracket, potentially increasing the demand for wild-sourced rhino horn.[9] Customers that are attracted by exclusivity and high price may directly source wild rhino horn. This could be achieved either through legal means such as trophy hunting, which is accompanied by CITES certification,[4] or illegally, through poaching. The predicted dynamics of a bio-fabricated rhino horn market need to be supported by statistical evidence.[7] It is currently impossible to accurately assess whether direct sourcing of wild rhino horn will take place. If Pembient's proposed model is to function, in order for bio-fabricated horn to flood the illicit market, prices must be kept low.[9] Companies would have to be legally obliged to sell at prices that best aid conservation efforts and not business

initiatives, which may be an unrealistic expectation.[9]

What Other Strategies Can Alleviate the Poaching Crisis?

Demand reduction is a crime reduction strategy that aims to reduce wildlife crime by decreasing poaching incentives.[19] This is a popular technique favoured amongst many conservationists.[20] Demand reduction techniques to combat rhino poaching include spreading scientific knowledge that rhino horn has no medicinal value and drawing public attention to rapidly declining rhino population numbers.[9] However, the use of rhino horn for medicinal purposes is engrained in culture and tradition, and has been endorsed by medical practitioners for centuries. An example being the 16th century pharmacist Li Shih Chen's 50-volume pharmacology text.[21] A change in outlook will take time. Legalising the trade of existing rhino horn stocks without the introduction of bio-fabricated horn is another option, however this presents a number of issues. Some buyers of rhino horn have shown clear preferences regarding whether lethal or non-lethal methods were used to harvest the horn they wish to purchase.[5] Legalising trade provides customers with horn from existing rhino horn stocks, while further animals may not be immediately poached, and the money generated can fund the protection of remaining rhinos.[20] However this may extend the market in which rhino horn is available, resulting in a demand increase.[20] Controlling this increased demand could exacerbate existing pressures on an already strained law enforcement.[16] Another issue is that the demand for wildlife products increases as human population increases, and eventually existing stocks will be depleted. Demand could then be satisfied through rhino farming.[20] The farming of wildlife products however is more expensive than poaching, an example being the farming of Tiger (Family *Felidae*) parts. It costs at least US $4000 to raise a tiger to adulthood, whereas poaching costs as little as $15–$20.[22] This creates an economic incentive for poaching in any legal market.[22] In addition to this issue, rhino horn farming raises animal welfare and rights issues.

Can Bio-fabricated Rhinoceros Horn Alleviate the Current Poaching Crisis?

The largest survey of rhino horn users in Vietnam consisted of a sample size of 900 out of which 857 responses were chosen for analysis.[5] This is the largest survey of rhino horn users in Vietnam. Our knowledge of the rhino horn market is inadequate. Larger sample sizes are required with questions that specifically address how bio-fabricated horn is viewed by the end-user. Existing data does not reveal whether bio-fabricated horn will be viewed a "fake" or an alternative, safer source of rhino horn.[7]

Conclusion

Preventing the Rhino's Extinction:

Future decisions concerning the sale of rhino horn need to be supported by a knowledge of consumer demand and accurate economic models of illicit and legal rhino horn markets. Proceeding without the support of further investigation could have numerous negative consequences. Loosening trade restrictions or tightening law enforcement could have unexpected effects on wildlife trade if the markets are not understood.[23] The sale of bio fabricated horn could increase demand for wild sourced horn and severely complicate law enforcement.[9] A further concern is that new technologies could be expensive and uneconomical and lastly that trade legalisation could intensify demand and lead to the rhino's extinction.[9,24]

How effective bio-fabricated substitute products are as a conservation tool is relatively unknown. Social, cultural and historical practices involving wildlife products have existed for thousands of years and will not cease instantly.[21] For most endangered species that are poached for wildlife products, an immediate solution is required. Bio-fabricated substitute products could offer an immediate solution for many wildlife trafficking issues, not just for rhinos.[7] Other Species such as Pangolins (Family *Manidae*)[25] and Elephants (Family *Elephantidae*)[26] could benefit. Bio fabricated pangolin scales and ivory could act as substitute products, relieving pressure on remaining populations and preventing extinctions.[7] Evidently research examining bio-fabricated substitute products is not just beneficial to rhino populations. It has universal implications for future conservation strategies.

To summarise it is clear that future decisions regarding the sale of rhino horn need to be supported by extensive further research. Bio-fabricated horn is only one potential method of alleviating poaching and prevent the looming extinction of the world's rhino populations.[7] Poaching is a significant threat that must be resolved in parallel to other issues such as habitat loss and fragmentation if successful conservation of the species is to be achieved.[6] If we fail to act quickly and appropriately, the rhinos' 60 million years of existence will abruptly end and the world will be deprived of these iconic species. South Africa will be home to the Big Four, the Big Five becoming an extinct natural phenomenon.[27]

References

1. Conservation on International Trade in Endangered Species of Wild Fauna and Flora. Appendices I, II and III. 1st ed. *CITES*; 2016. Available from: https://cites.org/sites/default/files/eng/app/2015/E-Appendices-2015-02-05.pdf

2. Hermes R, Göritz F, Portas TJ, Bryant BR, Kelly JM, Maclellan LJ, Keeley T, Schwarzenberger F, Walzer C, Schnorrenberg A, Spindler RE. Ovarian superstimulation, transrectal ultrasound-guided oocyte recovery, and IVF in rhinoceros. *Theriogenology*. 2009; 15;72(7):959-68.

3. Gao Y, Stoner K, Lee A, Clark S. Rhino horn trade in China: An analysis of the art

and antiques market. *Biological Conservation*. 2016; 201:343-347.

4. Milliken T. Illegal trade in ivory and rhino horn: An assessment report to improve law enforcement under the Wildlife TRAPS Project. *USAID and TRAFFIC*. 2014.

5. Hanley N, Sheremet O, Bozzola M, Kasterine A, MacMillan DC. Consumer Demand for Rhino Horn in Vietnam: insights from a choice experiment. Instruments and incentive mechanisms for biodiversity conservation and ecosystem service provision: *Biodiversity and Economics for Conservation; Cambridge* 2016. p. 36-7.
6. T. Milliken, J. Shaw, The South Africa–Viet Nam rhino horn trade nexus: A deadly combination of institutional lapses, corrupt wildlife industry professionals and Asian crime syndicates (*TRAFFIC, Johannesburg, 2012).*

7. Broad S Burgess G. Synthetic Biology, Product Substitution and the Battle Against Illegal Wildlife Trade. *Traffic Bulletin*, 2016;28(1):23

8. Cress D, Zommers Z. Emerging Technologies: Smarter ways to fight wildlife crime. *Environmental Development*. 2014; 12:62-72.

9. Centre for Biological Diversity *et al*., 2016. Before the Secretary of the Interior (USA): Petition to ban the import, export, and sale of "cultured" rhinoceros horn and cultured products of other protected wildlife species - February 10, 2016. *Washington, CBD -*, pp. 1-10.Available from: http://www.rhinoresourcecenter.com/index.

10. Hieronymus TL, Witmer LM, Ridgely RC. Structure of white rhinoceros (Ceratotherium simum) horn investigated by X-ray computed tomography and histology with implications for growth and external form. *Journal of Morphology*. 2006; 1;267(10):1172-6.

11. Bose S, Vahabzadeh S, Bandyopadhyay A. Bone tissue engineering using 3D printing. *Materials Today.* 2013; 16(12):496-504.

12. Markus M. Proposed Role of Bio Fabricated Horn in Addressing Illegal Wildlife Trade. *Pembient, Inc;* 2016. Available from: http://www.3dprintingbusiness.directory/wp-content/uploads/2016/08/AfRSG-Pembient-White-Paper-2016-DRAFT2.pdf?x98208

13. Akerlof G. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economic*s. 1970;84(3):488.

14. Save the Rhino. Synthetic Rhino Horn: Will it save the rhino? [Internet]. *Savetherhino.org*. 2015. Available from: https://www.savetherhino.org/rhino_info/thorny_issues/synthetic_rhino_horn_will_it_save_the_rhino

15. France-Presse A. The Guardian. Zulu false dawn: Shembe faithful swap leopard skin for faux fur. Available from: https://www.theguardian.com/world/2014/feb/19/zulu-shembe-leopardskin-south-africa.

16. Leader-Williams, NMilner-Gulland E. Policies for the Enforcement of Wildlife Laws: The Balance between Detection and Penalties in Luangwa Valley, Zambia. *Conservation Biology.* 1993;7(3):611-617.

17. CITES Secretariat. Interpretation and implementation of the Convention Species trade and conservation Rhinoceroses. Convention on the International Trade in Endangered Species of wild fauna and flora Sixteenth meeting of the Conference of the Parties Bangkok (Thailand). *Report: CoP16 Doc*. 54.2, 2013. Available from https://cites.org/eng/cop/16/doc/E-CoP16-54-02.pdf.

18. Ogden R, Dawnay N, McEwing R. Wildlife DNA forensics—bridging the gap between conservation genetics and law enforcement. *Endangered Species Research.* 2009; 179-195.

19. Schneider JL. Reducing the illicit trade in endangered wildlife: the market reduction approach. *Journal of Contemporary Criminal Justice*. 2008; 24(3):274-95.

20. t'Sas-Rolfes M. Rhino Poaching: What is

the Solution?. *Solutions*. 2016;7(1):38-45

21. Costa-Neto EM. Animal-based medicines: biological prospection and the sustainable use of zootherapeutic resources. *Anais da Academia Brasileira de Ciências*. 2005; 77(1):33-43.

22. Gratwicke B, Bennett E, Broad S, Christie S, Dutton A, Gabriel G, Kirkpatrick C, Nowell K. The world can't have wild tigers and eat them, too. *Conservation Biology*. 2008; 1:22(1):222-3.

23. Fischer C. The complex interactions of markets for endangered species products. *Journal of Environmental Economics and Management*. 2004; 48(2):926-953.

24. Crookes DJ, Blignaut JN. Debunking the myth that a legal trade will solve the rhino horn crisis: A system dynamics model for market demand. *Journal for Nature Conservation*. 2015 30;28:11-8.

25. Nijman V, Zhang MX, Shepherd CR. Pangolin trade in the Mong La wildlife market and the role of Myanmar in the smuggling of pangolins into China. *Global Ecology and Conservation*. 2016; 31;5:118-26.

26. Stiles D. The ivory trade and elephant conservation. *Environmental Conservation*. 2004; 1:31(04):309-21.

27. Menon V. Poaching and Protection of Greater One-horned Rhinoceroses in India. *Traffic International*. 1996 114pp. ISBN 1 85850 102 4. Available from: http://www.traffic.org/species-reports/traffic_species_mammals46.

Life Sciences

TS
SR

Letter from the Editors

The life sciences allows us to investigate the extraordinary complexity of living things on both a macroscopic and molecular level. We can now examine the molecular machinery of cells and catalogue their genetic codes, proteins and metabolic reactions. Research in the life sciences has provided clarity in a subject that was once prohibitively difficult. However, the more we learn about the astonishing complexity of life, the more we realise how much remains to be discovered.

This incomplete understanding is not just an academic annoyance, it impedes our ability to treat disease and improve human health. Advances in the life sciences provide a foundation for advances in the world of medicine. Research in the life sciences is also applied to global problems with biotechnology solutions to challenges in agriculture and climate change. The life sciences are therefore vital in improving human living standards in many ways. This what attracts many undergraduates to study the Life Sciences at Trinity.

In this year's edition of the Trinity Student Scientific Review, the papers published in the life sciences category describe advances in a variety of research areas. Some of these papers re-examine topics in which new evidence has challenged old assumptions, while others feature recent discoveries and novel disease therapies. Each of them however, discuss intriguing scientific progress and have earned their place in the journal.

The interest that students have for the Life Sciences is reflected in the volume of papers that are submitted to this category. As the number of submissions are higher for the life sciences than the other categories, this year a team of three editors was chosen, instead of a single editor as in the past. Even with this expanded editorial team, there was still a considerable amount of work involved in preparing this section.

This editorial process was guided by the comments and expert advice of our Peer Reviewers. The editors would like to thank our Peer Reviewers Killian Hanlon, Nicole Campbell, Dinorath Olvera Ramos, Jessica Walls, Elena Stavenschi, Proinnsias Fox, Karen Slattery and Natalie Adlesic for their opinions and advice which helped to improve these papers. As well as this, we would like to extend our sincere thanks to the life sciences Academic Advisor, Dr Aisling Dunne, as her expert opinions were essential in selecting and refining these papers.

Finally, we would like to congratulate all of the students who submitted papers to the life sciences this year, whether or not they were included in the final journal. This journal would not be possible without their

dedication and hard work. All of the papers which we received were well written, with clear descriptions of their research topics and informed perspectives on potential applications. It was a privilege for us to review these exemplary papers, we hope that readers will enjoy them as much as we have.

Jack Schofield, Conor McLoughlin & Kat Hughes

Life Sciences Editors
Trinity Student Scientific Review 2017

Immunotheraphy: at the Front Line in the Fight Against Cancer;

CTLA4 and PD-1 Checkpoint Inhibitors

Eoin Connor
Senior Sophister
Immunology

Our immune system has a remarkable ability to recognize and neutralize cancer cells by using a process known as immunosurveillance. The key immune cells involved in immunosurveillance are cytotoxic T lymphocytes and Natural Killer (NK) cells which act to detect and eliminate cancer or infected cells via the secretion of proteases and pro-inflammatory cytokines. However, cancer cells manage to evade immune detection by employing a number of complex mechanisms including the ability to actively suppress the activity of these effector cells. One mechanism through which immune evasion can be achieved is via "checkpoint proteins" such as Programmed cell Death protein-1 (PD-1) and Cytotoxic T-Lymphocyte Associated protein 4 (CTLA4) which are responsible for regulating inflammatory immune responses in order to prevent tissue damage or autoimmunity. Cancer cells "hijack" these immune checkpoint pathways in an effort to resist killing by the immune response of cytotoxic T cells and NK cells. Immunotherapy in cancer treatment involves enhancing the body's natural defences in order to fight cancer. Recent advances in immunotherapy have come as a result of the direct blockade of these immune checkpoints. Ipilumimab and Nivolumab are FDA approved monoclonal antibodies designed to block the interaction of CTLA4 and PD-1 with their respective ligands thus preventing the suppression of immune cells in the tumour microenvironment and allowing for the successful detection and elimination of cancer cells.

Introduction

The average human will produce around 10^{16} (ten quadrillion) cells throughout the course of their lifetime. Cells in highly proliferative tissues such as the skin, the gut endothelium or the cells of the immune system, which have lifespans ranging from hours to months, undergo relatively constant proliferation at a rate of 107 cells per second.[30]

With each cell division our DNA is duplicated by DNA polymerases and it is estimated that with each replication up to 1 million copy errors occur.[1] This, coupled with mutations arising from DNA damage as a result of environmental mutagens, creates huge potential for the occurrence of cancer *via* loss of function mutatiowns in genes that prevent cell proliferation and gain of function mutations in genes promoting proliferation. Cells are equipped with enzymes specific for the "proofreading" of DNA replication as well as DNA damage repair which reduce the risk of cancer-causing mutations from occurring. Some mutations do, however, escape these proofreading and repair enzymes, allowing the mutations to persist thus increasing the possibility of the cell becoming cancerous. 5-10% of cancers result from the inheritance of mutations from parents that increase an individual's susceptibility to developing cancer, this is known as inherited cancer.[35]

It is the role of our immune system, specifically the Cytotoxic T Lymphocytes ($CD8^+$ T cell), which are cells of the adaptive immune system and the NK cells, of the innate immune system, to detect these transformed cells and destroy them in a controlled manner, thus preventing the development of cancer.[31] These cytotoxic cells are responsible for patrolling the body and distinguishing "self" from "non-self" i.e. identifying, in the case of cancer, cells that display signs of mutation. This process of patrolling the body and guarding against cancer is known as "immunosurveillance".[2, 32]

In order for a tumour to survive and persist in the body it must evade immune detection and killing. This can be achieved *via* a number of mechanisms, the most common being the down-regulation of the Major Histocompatability Complex (MHC) class 1 which is present on all nucleated cells and is responsible for presenting self and non self peptides on the cell surface.[27] If a foreign antigen or tumour related protein is presented by MHC class 1, it is detected by the $CD8^+$ T cell and the cell is killed *via* the enzymes perforin, which perforates the cell membrane, and granzyme, which enters the cell and induces cell death by triggering various apoptotic pathways.[33,3] Natural Killer cells are also capable of killing cancer cells in a granzyme and perforin-dependent manner. NK cells detect stress signals on the surface of cancer cells which allows the NK cells to elicit their cytotoxic function and secrete proinflammatory cytokines. Cytokines are a broad family of small proteins that alter the behavior of surrounding cells *via* binding to surface receptors and are fundamental to the regulation of the immune response. These cytokines act on local macrophages and Dendritic Cells (DCs) which phagocytose, or "eat" the apoptotic, or dying, cancer cells, migrate to local draining lymph nodes

and present specific cancer proteins to naïve T lymphocytes i.e. T cells that have not been presented with a foreign peptide before and are therefore not specifically reactive against a protein.[36] These T cells become "activated" and respond specifically to tumour antigen presented on class I MHC.[37] Cancer cells can evade immune detection by down-regulating the expression of MHC class 1 on the cell surface, thus avoiding the detection of $CD8^+$ T cells. This, however, is not sufficient to escape immune detection by NK cells, which are capable of detecting cells that do not express MHC class 1 and effectively eliminate them. This is known as the "missing-self hypothesis".[4] NK cell function can be diminished within the tumour microenvironment as tumour cells sequester essential nutrients and oxygen from the innate immune cells thus inhibiting their effector function.[45]

Tumours are capable of subverting our immune system and progress to cause disease.[34] The manner in which cancer cells evade destruction is an area of active research in the field of immunology, and our increased understanding of cancer immune evasion strategies has allowed for the development of specific and effective anti-cancer therapies. Indeed, cancer immunotherapy and the examples to be discussed in this review, anti-CTLA4 and anti-PD-1 monoclonal antibodies, are proving to be one of the greatest successes in immunology of the 21st century.

Cytotoxic T Lymphocyte Associated Protein 4 (CTLA4)

In order for a T cell to become activated against a specific antigen, a foreign protein that elicits an immune response, three signals are required, provided by an Antigen Presenting Cell (APC), e.g. a DC or macrophage (Figure 1). The first signal is between the T cell receptor (TCR) and the antigen, presented in the context of MHC molecules. APCs are capable of phagocytosing apoptotic tumour cells and presenting tumour associated antigen (TAA) to T cells.[40,39] The second signal is the costimulatory signal provided by CD80/CD86, cell surface proteins expressed on APCs which binds to its receptor CD28, expressed on the cell surface of T cells, and is responsible for the proliferation, differentiation and survival of T cells.[41,6] Without this second signal, T cells enter a state of anergy where they are functionally inert and unresponsive.[42] Signal three involves the secretion of cytokines by the APC and determines the functional role of the now activated T cell.

In order to prevent autoimmune and autoinflammatory conditions, the activation of T cells must be tightly regulated. This is achieved in a number of ways, one of which is *via* co-inhibitory molecules such as CTLA4.[44] For their role in the regulation of the immune response, these co-inhibitory molecules are also referred to as immune checkpoint proteins.

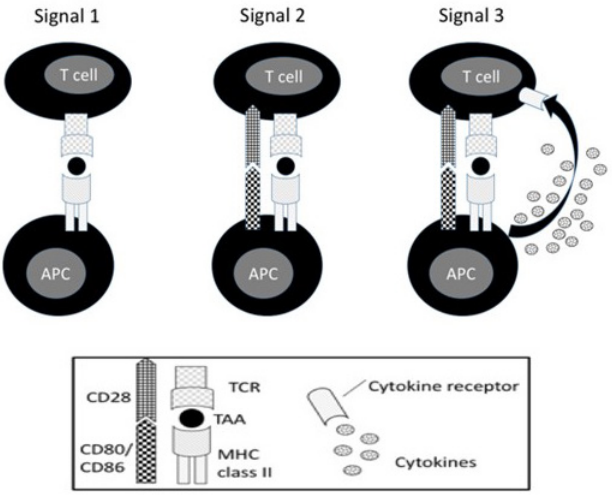


Figure 1: Three signals involved in T cell activation. Signal 1 is the interaction between the TCR and the antigen, presented as part of MHC class II molecules on APCs. Signal 2 comes from the interaction between the costimulatory molecules, CD80/CD86 on the APC and the CD28 coreceptor on the T cell. Signal 3 is the production and secretion of cytokine signalling molecules by the APC.
Adapted from Gutcher & Becher (2007).[43]

Twenty-four to forty-eight hours after T cell activation, CTLA4 is expressed on the T cell surface and migrates towards the immune synapse; the interface between the APC and T cell. CTLA4 actively competes for binding with the co-stimulatory molecules CD80/CD86 for which it has an affinity that is 20-40 times higher than that of CD28 thus inhibiting the costimulatory effect on the T cell.[8] The importance of CTLA4 as a negative regulator of T cells is highlighted by the fatal autoimmunity caused by a knock out mutation in the CTLA4 gene.[7]

Regulatory T cells (Tregs), the immune cells responsible for modulating the immune response and preventing immunopathology, constitutively express CTLA4. Recent evidence suggests that Tregs are capable of acquiring the CD80/CD86 co-stimulatory molecules from the surface of APCs *via* CTLA4-CD80/86 interactions, thus limiting the availability of co-stimulatory molecules that are capable of stimulating a T cell effector response.[9] CTLA4 downstream signaling also inhibits T cells from producing the major cytokine involved in T cell proliferation; Interleukin-2 (IL-2). Inhibition of this cytokine production, which is a stimulator of T cell growth, induces cell cycle arrest, thus inhibiting T cell proliferation.

Reverse signaling *via* CTLA4-CD80/86 may also induce the production of the immunoregulatory enzyme indoleamine 2,3-dioxygenase (IDO) by DCs which suppress T cell responses and clonal expansion.[10]

The presentation of tumour associated antigens (TAA) by APCs in the presence of CTLA4 renders the adaptive immune response anergic to the tumour thus allowing the tumour to remain undetected and proliferate i.e. the effector T cells become tolerant to the tumour.[46] With this discovery came the hypothesis that a blockade against CTLA4 may break this tolerance and allow effective clearance of the tumour.

Anti-CTLA4 Antibodies as Immunotherapy

Cancer Immunotherapy is the process of enhancing the immune system's natural capability to eliminate cancer and other diseases. The discovery of the immune-inhibitory effect of CTLA4 immediately led to the development of CTLA4-blocking antibodies which prevent the interaction between CTLA4 and CD80/86, also referred to as B7 (Figure 2).

The results from studies in murine models of human cancers, such as breast and prostate cancer, showed great potential for clinical efficacy in humans. Early trials in mice that were treated with a transplanted carcinoma found that the administration of anti-CTLA4 antibodies resulted in complete clearance of the carcinoma and immune memory was generated, thereby providing long lasting protection against subsequent transplanted tumours. This was first discovered in 1996 by researchers in the Berkeley Cancer Research Laboratory in California.[11]

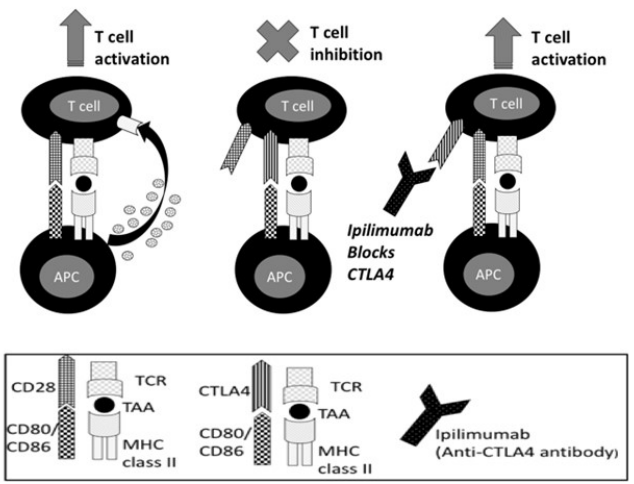


Figure 2: Illustration of the role of CTLA4 as an immune regulator. T cells are activated by the interaction between CD80/CD86 on the APC and CD28 on the T cell. CTLA4 inhibits this T cell activation. The CTLA4 antibody, Ipilimumab prevents this inhibition by CTLA4, which allows T cell activation.
Adapted from Saijo (2012).[15]

The CTLA4 blockade alone was not as effective at combating less immunogenic mice tumour models i.e. tumours that are less able to elicit an immune response. Co-operation between the innate and adaptive immune system is essential for effective tumour or pathogen clearance. If a tumour is incapable of activating the innate immune system due to a lack of immunogenicity, the innate cells cannot activate effector T cells and the tumour will persist.

Studies carried out by Elsas and colleagues using B16 melanoma in mouse models demonstrated that when the CTLA4 blockade was combined with Granulocyte/Monocyte Stimulating Factor (GMCSF)-producing vaccines, which act by inducing proliferation and accumulation of DCs at the site of injection, they observed tumour eradication in 80% of B16 melanoma mouse models (Figure 3).[12,47] Similar studies also showed a significant improvement in survival rates when CTLA4 treatment was combined with vaccines comprised of tumour antigens which act by "priming" the innate cells allowing for effective induction of an adaptive T cell response. The success in murine models led to the development of Ipilimumab, the human anti-CTLA4 antibody, by researchers in the University of California, Berkeley.[14]

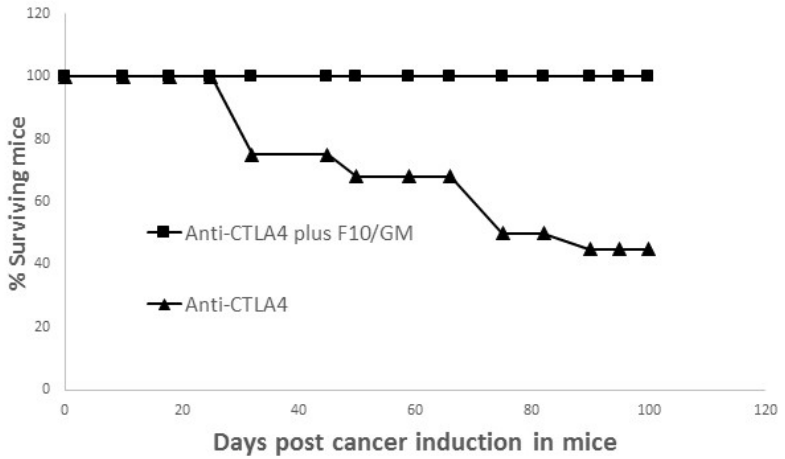


Figure 3: The results of Anti-CTLA4 antibody plus GMCSF-producing vaccine (F10/GM) combination therapy in murine B16 melanoma.
Adapted from Allison, van Elsas & Hurwitz (1999).[12]

Ipilimumab was granted FDA approval for the treatment of metastatic melanoma in 2011. Initial studies showed that combination treatment of Ipilimumab with a cancer antigen based vaccine increased patient survival by 10 months when compared to untreated controls. In a clinical trial from Bristol-Myers Squibb ipilimumab was combined with Dacarbazine (DTIC), a widely used chemotherapeutic. Patients receiving Ipilimumab had a mean survival of 11.2 months and patients receiving DTIC alone had a mean survival of just 9 months.[13]

As mentioned previously, mutations in CTLA4 can result in severe autoimmune and autoinflammatory conditions. It was therefore unsurprising when patients undergoing Ipilimumab treatment displayed symptoms of autoimmune inflammation such as dermatitis, nephritis and pancreatitis, and in one trial 14 deaths were caused as a result of Ipilimumab treatment.[14] The objective of checkpoint inhibitors such as Ipilimumab, is to break T cell tolerance to tumours and in doing so, we effectively inhibit T cell inhibition. While this is beneficial with regards to cancer treatment, immune checkpoints act to maintain homeostasis in a healthy individual, preventing uncontrolled inflammation and immunopathology. Blocking these immune checkpoints disturbs this natural homeostasis and can result in Immune-related adverse effects (IRAE) which are often severe requiring further medication in order to suppress the uncontrolled immune response.[48] The quality of life of patients must be considered when discussing immunotherapeutics. Does the benefit of potentially increasing the patient's chances of survival by 10 months out-weigh the risk of the patient developing a severe and life-altering autoimmune disease?

Programmed Cell Death-1 (PD-1)

As discussed previously, the immune response must be tightly regulated such that any threat to the body is eliminated while maintaining homeostasis and limiting immunopathology. PD-1, in addition to CTLA4, is an immune checkpoint expressed on T/B lymphocytes and is associated with immune exhaustion.[18] PD-1 is highly expressed on T cells involved in chronic infection, and is responsible for shutting down an immune response by inducing T cell apoptosis, a form of controlled cell death.[49]

The function of PD-1 was determined through knock out studies in mice, revealing that, unlike CTLA4 which when mutated results in dramatic systemic autoimmunity, PD-1 deficient mice exhibit organ specific autoimmunity such as lupus, autoimmune diabetes and hepatitis due to the high expression levels of PD-L1, the ligand for PD-1, on endothelial cells in sites commonly associated with such organ specific autoimmune disorders.[16] Knock out studies also demonstrated the pivotal role PD-1 plays in maintaining immune tolerance during chronic viral infection in which PD-1 deficient mice develop fatal autoimmunity.[50] Similar to CTLA4, PD-1 expression on the T cell surface is induced upon lymphocyte

activation and expression peaks 48 hours after stimulation.[5]

The ligands for PD-1 are PD-L1 and PD-L2. PD-L1 is commonly expressed on macrophages and DCs as well as endothelial and epithelial cells. PD-L1 expression is induced by proinflammatory cytokines such as Interferon-γ (IFN-γ) and Tumour necrosis factor alpha (TNFα). PD-L1 deficiencies have been associated with exacerbated T cell responses and organ specific autoimmunity.[17] The role of PD-L2 is less clear, however its expression is restricted to immune cells.

It has also been shown that PD-1-PD-L1 interactions between a naïve T cell and an APC regulate the development and function of Tregs by enhancing and sustaining the expression of the key Treg master regulator Forkhead box p3 (FOXP3)[19]. This discovery suggests that tumour cells are capable of using the PD-1-PD-L1 pathway in order to subvert the immune system and evade detection by increasing the ratio of Tregs to effector T cells in the tumour microenvironment which becomes increasingly immunosuppressive. Indeed, it was discovered that PD-L1 expression by tumour cells correlates with poor survival.[38,51] The same study from Zhang *et al.* and colleagues demonstrated that the expression of PD-1 on tumour infiltrating T lymphocytes also correlates with a poor prognosis.[20]

These findings prompted the development of a PD-1/PD-L1 blockade anti-cancer immunotherapy.

PD-1/Pd-L1 Blockade as Immunotherapy in Cancer

Given the evidence supporting the role of PD-1-PD-L1 interactions in creating an immunosuppressive tumour microenvironment and allowing a tumour to subvert an anti-tumour immune response, huge effort was invested to develop a PD-1/PD-L1 blocking antibody. Early murine studies showed great promise, with antibodies against both PD-1 and PD-L1 inhibiting the spread of B16 melanoma and colon cancer *via* $CD8^+$ effector T cells.[12] These studies also increased our understanding of the mechanisms that occur in the creation of an immunosuppressive tumour microenvironment. It was found that IFN, produced by $CD8^+$ T cells during an effector response, causes the tumour cells to up-regulate the surface their expression of PD-L1. This was shown to be reversed *via* the use of anti-IFN neutralizing antibodies.[26]

The first antibody against PD-1 to be approved by the FDA was Nivolumab in 2014, for the treatment of metastatic melanoma, and in 2015, for the treatment of metastatic squamous non-small cell lung cancer. Pembrolizumab was also approved in 2014 for patients with advanced melanoma and in 2016 for the treatment of metastatic head and neck squamous cell carcinoma.[21,22,23]

Recent studies have compared the effect of Nivolumab versus Dacarbazine, an approved chemotherapeutic drug, against metastatic melanoma. The results

showed that after one year the survival rate for the group given Dacarbazine was 42.1% compared to a 72.9% survival rate for the Nivolumab group. On average, Nivolumab increased survival by 8.9 months whereas Dacarbazine increased survival by 6.8 months.[24] Nivolumab was shown to carry out its effect by binding to PD-1 on tumour infiltrating CD8+ T cells thus inhibiting the interaction between PD-1 and its ligand PD-L1. [24] A similar study carried out by Reck *et. al* comparing Pembrolizumab to existing platinum-based chemotherapy for Non-Small-Cell Lung Cancer revealed a survival rate after 6 months of 80.2% in the Pembrolizumab treated group compared to 72.4% in the chemotherapy group.[25]

Discussion

The recent breakthroughs in cancer treatment are based on immunotherapeutics, with checkpoint inhibitors such as anti-CTLA4 and anti-PD-1 monoclonal antibodies taking center stage in the fight against cancer. The development of these immunotherapeutics has not come from efforts to improve the existing chemotherapeutics and radiation therapies, but rather has come as a result of increasing our understanding of tumour immunology. We have returned to the drawing board with regards to cancer treatment and this has provided us with fruitful results. By furthering our understanding of how a tumour subverts the immune system and uncovering the mysteries of the tumour micro-environment, there is no doubt that the area of immunotherapy will continue to expand and produce successful results.

The use of checkpoint inhibitors in cancer treatment does not come without risk. Immune-related adverse effects (IRAE) occurred in 24% of patients receiving anti PD-1 antibody treatment compared to 85% of patients receiving anti-CTLA4 treatment. The most common IRAEs resulting from anti-PD1 treatment were systemic lupus erythematosus and Rheumatoid Arthritis.[53] Treatment with immunosuppressive drugs proved effective in treating the IRAEs without compromising the anti cancer treatment. However, long term treatment with immunosuppressive drugs does leave patients susceptible to opportunistic infections as well as cancer.[54] The quality of life of patients receiving cancer treatment must be considered. Is it ethical to introduce treatments that have a high probability of causing severe and life-altering autoimmune-like symptoms in order to increase survival rate by a number of months?

Controversy also surrounds the lack of efficacy of checkpoint inhibitors as a monotherapy. As shown by the study carried out by Elsas and colleagues, anti-CTLA4 antibodies had minimal effects of less immunogenic tumours and only showed clinical significance when combined with existing chemotherapeutics.[12] Similar results were also shown for the use of anti-PD-1 as a monotherapy.[29] While the development of these checkpoint inhibitors is an important step in cancer treatment we are still reliant on existing chemotherapies and radiation therapies, which also cause severe side effects, to use in combination with these treatments.

While the future of immunotherapy does appear bright there is still controversy surrounding its use in cancer treatment. The objectives of researchers should be to replace the existing chemotherapeutics and radiation therapies with highly selective anti-cancer therapies that specifically target tumour cells with minimal adverse effects. Combination therapies of Nivolumab plus Ipilimumab versus a monotherapy of ipilimumab in melanoma patients revealed a significant overall increase in complete response rates in patients receiving the combination therapy. Nivolumab in combination with Ipilimumab was granted FDA approval for the treatment of metastatic melanoma in 2015 despite 45% of patients receiving the combination therapy withdrawing from the trial due to severe adverse effects.[52]

References

1. Preston BD, Albertson TM, Herr AJ. DNA replication fidelity and cancer. 2010 Oct 15 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2993855/.

2. Gonzalez S, González-Rodríguez AP, Suárez-Álvarez B, López-Soto A, Huergo-Zapico L, Lopez-Larrea C. Conceptual aspects of self and nonself discrimination. 2011 Jan 1 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136900/.

3. Hewitt EW. The MHC class I antigen presentation pathway: Strategies for viral immune evasion. 2003 Oct 1;110(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1783040/.

4. Fruci D, Benevolo M, Cifaldi L, Lorenzi S, Lo Monaco E, Tremante E, Giacomini P. Major histocompatibility complex class i and tumour immuno-evasion: How to fool T cells and natural killer cells at one time. 2012 Feb 1. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3267590/.

5. Jin H, Ahmed R, Okazaki T. Role of PD-1 in regulating t-cell immunity. Current topics in microbiology and immunology. 2010 Nov 10;350:17–37. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21061197.

6. Fuse S, Obar JJ, Bellfy S, Leung EK, Zhang W, Usherwood EJ. CD80 and CD86 control antiviral $CD8^+$ t-cell function and immune surveillance of Murine Gammaherpesvirus 68. 2006 Sep 1;80(18). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1563936/.

7. Schubert D, Bode C, Kenefeck R, Hou T, Wing J, Kennedy A, Bulashevska A, Petersen B, Schäffer A, Grüning B, Unger S, Frede N, Baumann U, Witte T, Schmidt R, Dueckers G, Niehues T, Seneviratne S, Kanariou M, Speckmann C, Ehl S, Rensing-Ehl A, Warnatz K, Rakhmanov M, Thimme R, Hasselblatt P, Emmerich F, Cathomen T, Backofen R, Fisch P, Seidl M, May A, Schmitt-Graeff A, Ikemizu S, Salzer U, Franke A, Sakaguchi S, Walker L, Sansom D, Grimbacher B. Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. *Nature medicine*. 2014 Oct 21;20(12):1410–6. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25329329.

8. Mocellin S, Nitti D. CTLA-4 blockade and the renaissance of cancer immunotherapy. Biochimica et Biophysica Acta (BBA) - *Reviews on Cancer*. 2013 Dec 1;1836(2):187–96. Available from: http://www.sciencedirect.com/science/article/pii/S0304419X13000358 doi: 10.1016/j.bbcan.2013.05.003.

9. Qureshi O, Zheng Y, Nakamura K, Attridge K, Manzotti C, Schmidt E, Baker J, Jeffery L, Kaur S, Briggs Z, Hou T, Futter C, Anderson G, Walker L, Sansom D. Trans-endocytosis of CD80 and CD86: A molecular basis for the cell-extrinsic function of CTLA-4. *Science (New York, N.Y.)*. 2011 Apr 9;332(6029):600–3. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21474713.

10. Intlekofer A, Thompson C. At the bench: Preclinical rationale for CTLA-4 and PD-1 blockade as cancer immunotherapy. *Journal of leukocyte biology*. 2013 Apr 30;94(1):25–39. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23625198.

11. Leach, Krummel M, Allison J. Enhancement of antitumour immunity by CTLA-4 blockade. *Science (New York, N.Y.)*. 1996 Mar 22;271(5256):1734–6. Available from: https://www.ncbi.nlm.nih.gov/pubmed/8596936/.

12. van Elsas A, Hurwitz AA, Allison JP. Combination Immunotherapy of B16 Melanoma using Anti–Cytotoxic T Lymphocyte–Associated antigen 4 (Ctla-4) and Granulocyte/Macrophage colony-stimulating factor (Gm-Csf)-producing vaccines induces rejection of Subcutaneous and Metastatic tumours accompanied . 1999 Aug 2;190(3). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2195583/.

13. Wolchok JD, Hodi SF, Weber JS, Allison JP, Urba WJ, Robert C, O'Day SJ, Hoos A, Humphrey R, Berman DM, Lonberg N, Korman AJ. Development of ipilimumab: A novel immunotherapeutic approach for the treatment of advanced melanoma. 2013 Jun 17;1291(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3910157/.

14. Hodi F, O'Day S, McDermott D, Weber R, Sosman J, Haanen J, Gonzalez R, Robert C, Schadendorf D, Hassel J, Akerley W, den van, Lutzky J, Lorigan P, Vaubel J, Linette G, Hogg D, Ottensmeier C, Lebbé C, Peschel C, Quirt I, Clark J, Wolchok J, Weber J, Tian J, Yellin M, Nichol G, Hoos A, Urba W. Improved survival with ipilimumab in patients with metastatic melanoma. *The New England Journal of Medicine*. 2010 Jun 8;363(8):711–23. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20525992/.

15. Saijo N. Present status and problems on molecular targeted therapy of cancer. Cancer research and treatment : *Official Journal of Korean Cancer Association*. 2012 Apr 14;44(1):1–10. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22500155.

16. Yao S, Chen L. PD-1 as an immune Modulatory receptor. ;20(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4455017/.

17. Eppihimer M, Gunn J, Freeman G, Greenfield E, Chernova T, Erickson J, Leonard J. Expression and regulation of the PD-L1 immunoinhibitory molecule on microvascular endothelial cells. *Microcirculation* (New York, N.Y. : 1994). 2002 Apr 5;9(2):133–45. Available from: https://www.ncbi.nlm.nih.gov/pubmed/11932780/.

18. Riella LV, Paterson AM, Sharpe AH, Chandraker A. Role of the PD-1 pathway in the immune response. 2012 Aug 17;12(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3784243/#R77.

19. Francisco L, Salinas V, Brown K, Vanguri V, Freeman G, Kuchroo V, Sharpe A. PD-L1 regulates the development, maintenance, and function of induced regulatory T cells. *The Journal of Experimental Medicine*. 2009 Dec 17;206(13):3015–29. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20008522.

20. Zhang Y, Kang S, Shen J, He J, Jiang L, Wang W, Guo Z, Peng G, Chen G, Liang W. Prognostic significance of programmed cell death 1 (PD-1) or PD-1 ligand 1 (PD-L1) expression in Epithelial-Originated cancer. 2015 Feb 13;94(6). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4602735/.

21. Sul J, Blumenthal G, Jiang X, He K, Keegan P, Pazdur R. FDA approval summary: Pembrolizumab for the treatment of patients with Metastatic non-small cell lung cancer whose tumours express programmed death-ligand 1. *The Oncologist*. 2016 Mar 31;21(5):643–50. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27026676.

22. Poole R. Pembrolizumab: First global approval. *Drugs*. 2014 Oct 22;74(16):1973–81. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25331768.

23. Kazandjian D, Suzman D, Blumenthal G, Mushti S, He K, Libeg M, Keegan P, Pazdur R. FDA approval summary: Nivolumab for the treatment of Metastatic non-small

cell lung cancer with progression on or after platinum-based chemotherapy. *The Oncologist*. 2016 Mar 18;21(5):634–42. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26984449.

24. Robert C, Long G, Brady B, Dutriaux C, Maio M, Mortier L, Hassel J, Rutkowski P, McNeil C, Kalinka-Warzocha E, Savage K, Hernberg M, Lebbé C, Charles J, Mihalcioiu C, Chiarion-Sileni V, Mauch C, Cognetti F, Arance A, Schmidt H, Schadendorf D, Gogas H, Lundgren-Eriksson L, Horak C, Sharkey B, Waxman I, Atkinson V, Ascierto P. Nivolumab in previously untreated melanoma without BRAF mutation. *The New England Journal of Medicine.* 2014 Nov 18;372(4):320–30. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25399552.

25. Reck M, Rodríguez-Abreu D, Robinson A, Hui R, Csőszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A, Cuffe S, O'Brien M, Rao S, Hotta K, Leiby M, Lubiniecki G, Shentu Y, Rangwala R, Brahmer J, Investigators K. Pembrolizumab versus chemotherapy for PD-L1-Positive non-small-cell lung cancer. *The New England Journal of Medicine.* 2016 Oct 11;375(19):1823–33. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27718847?dopt=Abstract.

26. Chen L, Han X. Anti-PD-1/PD-L1 therapy of human cancer: Past, present, and future. *The Journal of Clinical Investigation*. 2015 Sep 2;125(9):3384–91. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26325035.

27. Ishido S, Wang C, Lee B-S, Cohen GB, Jung JU. Downregulation of Major Histocompatibility complex class I molecules by Kaposi's sarcoma-associated Herpesvirus K3 and K5 proteins. 2000 Jun 1;74(11). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC110885/.

28. Wolchok J, Kluger H, Callahan M, Postow M, Rizvi N, Lesokhin A, Segal N, Ariyan C, Gordon R, Reed K, Burke M, Caldwell A, Kronenberg S, Agunwamba B, Zhang X, Lowy I, Inzunza H, Feely W, Horak C, Hong Q, Korman A, Wigginton J, Gupta A, Sznol M. Nivolumab plus ipilimumab in advanced melanoma. *The New England Journal of Medicine*. 2013 Jun 4;369(2):122–33. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23724867.

29Sundar R, Cho B-C, Brahmer JR, Soo RA. Nivolumab in NSCLC: Latest evidence and clinical potential. ;7(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4346216/.

30. Bansbach CE, Cortez D. Defining genome maintenance pathways using functional Genomic approaches. ;46(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3144553/.

31. Arum C-J, Anderssen E, Viset T, Kodama Y, Lundgren S, Chen D, Zhao C-M. Cancer Immunoediting from Immunosurveillance to tumour escape in Microvillus-Formed niche: A study of Syngeneic Orthotopic rat bladder cancer model in comparison with human bladder cancer1. 2010 Jun 1;12(6). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2887496/.

32. Ljunggren H, Kärre K. In search of the 'missing self': MHC molecules and NK cell recognition. *Immunology Today.* 1990 Jul 1;11(7):237–44. Available from: https://www.ncbi.nlm.nih.gov/pubmed/2201309.

33. Smyth M, Trapani J. Granzymes: Exogenous proteinases that induce target cell apoptosis. *Immunology Today.* 1995 Apr 1;16(4):202–6. Available from: https://www.ncbi.nlm.nih.gov/pubmed/7734049.

34. Wong A, Mankovitz R, Kennedy J. Immunosuppressive and immunostimulatory factors produced by malignant cells in vitro. *International Journal of Cancer.* 1974 Apr 15;13(4):530–42. Available from: https://www.ncbi.nlm.nih.gov/pubmed/4601698.

35. Knudson AG. Mutation and cancer: Statistical study of retinoblastoma. 1971 Apr 1;68(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC389051/.

36. Thomas A, Santarsiero L, Lutz E, Armstrong T, Chen Y, Huang L, Laheru D, Goggins M, Hruban R, Jaffee E. Mesothelin-specific CD8(+) T cell responses provide evidence of in vivo cross-priming by antigen-presenting cells in vaccinated

pancreatic cancer patients. *The Journal of Experimental Medicine.* 2004 Aug 4;200(3):297–306. Available from: https://www.ncbi.nlm.nih.gov/pubmed/15289501.

37. Lee P, Yee C, Savage P, Fong L, Brockstedt D, Weber J, Johnson D, Swetter S, Thompson J, Greenberg P, Roederer M, Davis M. Characterization of circulating T cells specific for tumour-associated antigens in melanoma patients. *Nature Medicine*. 1999 Jun 17;5(6):677–85. Available from: https://www.ncbi.nlm.nih.gov/pubmed/10371507.

38. Cedrés S, Ponce-Aix S, Zugazagoitia J, Sansano I, Enguita A, Navarro-Mendivil A, Martinez-Marti A, Martinez P, Felip E. Analysis of expression of programmed cell death 1 ligand 1 (PD-L1) in malignant pleural mesothelioma (MPM). 2015 Mar 16;10(3). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4361537/.

39. Henry F, Bretaudeau L, Barbieux I, Meflah K, Gregoire M. Induction of antigen presentation by macrophages after phagocytosis of tumour apoptotic cells. *Research in Immunology.* 1998 Dec 16;149:673–9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/9851522.

40. Escors D. Tumour immunogenicity, antigen presentation and immunological barriers in cancer immunotherapy. *New Journal of Science.* 2014 Mar 19;2014. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24634791.

41. Gimmi C, Freeman G, Gribben J, Sugita K, Freedman A, Morimoto C, Nadler L. B-cell surface antigen B7 provides a costimulatory signal that induces T cells to proliferate and secrete interleukin 2. *Proceedings of the National Academy of Sciences of the United States of America.* 1991 Aug 1 [;88(15):6575–9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/1650475/.

42. Appleman L, Boussiotis V. T cell anergy and costimulation. Immunological reviews. 2003 Apr 3;192:161–80. Available from: https://www.ncbi.nlm.nih.gov/pubmed/12670403.

43. Gutcher I, Becher B. APC-derived cytokines and T cell polarization in autoimmune inflammation. 2007 May 1;117(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1857272/.

44. Brunet J, Denizot F, Luciani M, Roux-Dosseto M, Suzan M, Mattei M, Golstein P. A new member of the immunoglobulin superfamily--cTLA-4. *Nature.* 1987 Jul 16;328(6127):267–70. Available from: https://www.ncbi.nlm.nih.gov/pubmed/3496540?access_num=3496540&link_type=MED&dopt=Abstract.

45. Baginska J, Viry E, Paggetti J, Medves S, Berchem G, Moussay E, Janji B. The critical role of the tumour Microenvironment in shaping natural killer cell-mediated anti-tumour immunity. 2013 Dec 25;4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3872331/.

46. Staveley-O'Carroll K, Sotomayor E, Montgomery J, Borrello I, Hwang L, Fein S, Pardoll D, Levitsky H. Induction of antigen-specific T cell anergy: An early event in the course of tumour progression. *Proceedings of the National Academy of Sciences of the United States of America.* 1998 Mar 14;95(3):1178–83. Available from: https://www.ncbi.nlm.nih.gov/pubmed/9448305.

47. Gupta R, Emens LA. GM-CSF-Secreting vaccines for solid tumours: Moving forward.;10(50). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3086372/.

48. Bertrand A, Kostine M, Barnetche T, Truchetet M-E, Schaeverbeke T. Immune related adverse events associated with anti-cTLA-4 antibodies: Systematic review and meta-analysis. 2015 Sep 4;13. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4559965/.

49. Ishida Y, Agata Y, Shibahara K, Honjo T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *The EMBO Journal*. 1992 Nov 1;11(11):3887–95. Available from: https://www.ncbi.nlm.nih.gov/pubmed/1396582.

50. Barber D, Wherry E, Masopust D, Zhu B, Allison J, Sharpe A, Freeman G, Ahmed R. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature*. 2005 Dec 31;439(7077):682–7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/16382236.

51. Qing Y, Li Q, Ren T, Xia W, Peng Y, Liu G, Luo H, Yang Y, Dai X, Zhou S, Wang D. Upregulation of PD-L1 and APE1 is associated with tumourigenesis and poor prognosis of gastric cancer. *Drug design, Development and Therapy*. 2015 Mar 4;9:901–9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25733810.

52. Postow MA, Chesney J, Pavlick AC, Robert C, Grossmann K, McDermott D, Linette GP, Meyer N, Giguere JK, Agarwala SS, Shaheen M, Ernstoff MS, Minor D, Salama AK, Taylor M, Ott PA, Rollin LM, Horak C, Gagnier P, Wolchok JD, Hodi FS. Nivolumab and Ipilimumab versus Ipilimumab in untreated Melanoma. *New England Journal of Medicine*. 2015 May 21;372(21):2006–17.

53. Marrone K, Ying W, Naidoo J. Immune-related adverse events from immune Checkpoint inhibitors. *Clinical Pharmacology and Therapeutics*. 2016 May 14;100(3):242–51. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27170616.

54. Friedman C, Proverbs-Singh T, Postow M. Treatment of the immune-related adverse effects of immune Checkpoint inhibitors: A review. *JAMA Oncology*. 2016 Jul 2;2(10):1346–53. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27367787.

The Role of HIF-1α in the Re-education of Tumour Associated Macrophages

Hannah Prendeville
Senior Sophister
Human Health and Disease

The role of a cell's metabolic profile in sculpting its effector function, has become a major research interest in the last five years. With respect to tumour biology, it is becoming increasingly obvious that the metabolic profile of a tumour cell not only affects its own fate, but too the surrounding cells in the tumour micro-environment (TME). Tumour cells undergo a metabolic shift resulting in increased glucose consumption and flux through glycolysis, termed aerobic glycolysis. This shift re-wires the cell's metabolism to promote biosynthesis rather than energy production, to support rapid tumour cell proliferation. Tumour-derived lactic acid, a by-product of aerobic glycolysis, induces the functional polarisation of macrophages in the TME to promote a pro-tumourigenic, anti-inflammatory phenotype. This re-education occurs via stabilisation and translocation of hypoxia-inducible factor 1α (HIF-1α) to increase gene expression of vascular endothelial growth factor (VEGF) and arginase-1 (arg-1) which encourage tumour growth and metastasis. Alongside a PKM2 dimer, HIF-1α can also induce a pro-inflammatory reponse in activated macrophages, by upregulating the expression of IL-1β, a pro-inflammatory cytokine. It is essential to understand the different mechanisms used to induce each HIF-1α-driven macrophage phenotype in order to exploit these processes clinically. HIF-1α serves as a potential therapeutic target in cancer immunotherapy as it may be manipulated to promote a pro-inflammatory, anti-tumour response, similar to that seen during macrophage activation.

Introduction

Macrophages are a fundamental, plastic component of the innate immune system which perform a variety of functions ranging from establishing an inflammatory response, to tissue remodeling and wound repair.[1] The two broad macrophage categories include the classically activated, pro-inflammatory subset (Classically activated macrophages- CAMs) and the alternativaly activated, anti-inflammatory subset (alternatively activated macrophages- AAMs). Each set is generated in response to different environmental cues, including lipopolysaccharide (LPS) and IL-4 respectively, and display different metabolic profiles (Figure 1).[2,3] CAMs are characterised by increased expression of the inducible form of nitric oxide synthase (iNOS) and production of reactive oxygen species (ROS) and pro-inflammatory cytokines, thus play a role in the clearance of microbial infections. They utilise aerobic glycolysis for ATP production, resulting in the accumulation of metabolic intermediates to encourage a pro-inflammatory response. Conversely, AAMs are upregulated at the end of an infection and are known to regulate inflammation and promote wound healing. AAMs use the more efficient oxidative phosphorylation (OXPHOS) as a means of energy production.[4]

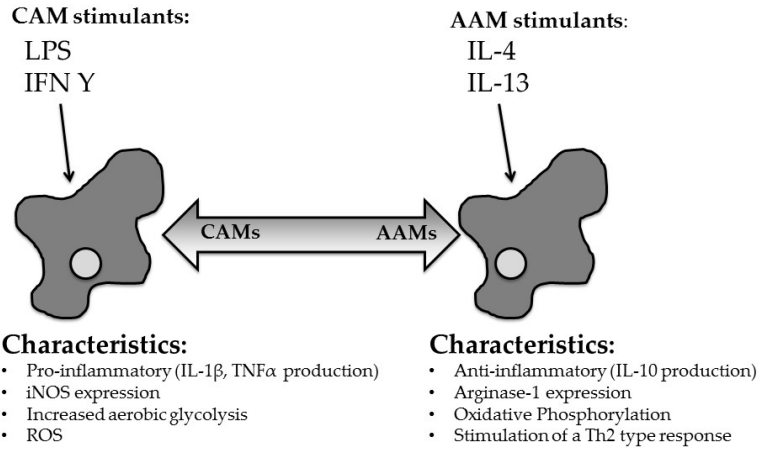


Figure 1: Characteristics of the major macrophage subsets. Both macrophage cateogries are functionally and metabolically distinct. CAMs are associated with inducing an inflammatory reponse through the production of pro-inflammatory cytokines such as IL-1β, and ROS. They use aerobic glycolysis as a method of ATP production, in order to enhance their pro-inflammatory phenotype. AAMs are primarily anti-inflammatory and promote wound healing and tissue repair, too stimulating a Th2 response. AAMs favor ATP production by OXPHOS. *This figure has been adapted with permission from Clària et al. (2011).*[5]

To carry out their primary task of maintaining tissue homeostasis, macrophages constantly monitor the functional state of tissue parenchymal cells and respond accordingly.[6] Similarly, tumour associated macrophages (TAMs) present in the

TME of solid tumours, also perform key homeostatic functions that promote tumour growth and progression. A TME is an intercellular communication network established by intricate interactions between malignant and non-transformed cells.[7] The non-malignant cell population includes a large proportion of infiltrating leukocytes, of which macrophages comprise the dominant portion. The effector functions of TAMs are pro-tumoural, reflected in their anti-inflammatory phenotype, similar to that of AAMs and their presence is associated clinically with a poor prognosis.[8]

Macrophages are highly plastic cells and their environment plays a major role in defining their phenotype and subsequent effector function. It is understood that macrophage re-education occurs in the TME where a switch from a tumour-suppressing CAM phenotype to a tumour-promoting AAM-like profile occurs.[9] It was only recently uncovered that tumour-derived lactic acid, a by-product of aerobic glycolysis, plays a critical role in driving the pro-tumourigenic, TAM phenotype, by inducing VEGF and arg-1 expression.[10] The exact signalling mechanisms promoting this phenotypic change aren't entirely understood. As the band of clinical research coined "immunotherapy", which aims to exploit the host's immune system to fight cancer, is becoming an increasingly attractive option for cancer treatment, understanding the mechanisms behind macrophage re-education by tumours is essential to exploit this process clinically.[11] Macrophages constitute the body's first line of defense, thus are an intelligent tool with the potential to be manipulated in order to promote an anti-tumour response.

This review will discuss the link between tumour cell metabolism and macrophage re-education and subsequent tumour growth, propose a potential molecule which may be key to driving the TAM phenotype and finally, explain the emerging importance of immunotherapy with respect to TAMs.

Tumour Cell Metabolism and its Link to Cellular Proliferation.

Otto Waburg was the first to characterise the metabolic profile of tumour cells when he observed elevated rates of glucose uptake, reduced rates of OXHPOS and persistant lactate production in the presence of oxygen. He coined this phenomenon "aerobic glycolysis".[12] It is now understood that this metabolic phenotype supports anabolic processes such as nucleotide synthesis *via* the pentose phosphate pathway (PPP), by channelling carbon intermediates for macromolecule production. Biosynthesis is thus enhanced, contributing to an environment amenable for cell proliferation, supporting tumour growth.[13]

Approximately 75 years following Warburg's discovery, the Cantley lab showed that this switch in cellular metabolism relies on the expression of one particular isoform of the glycolytic enzyme, pyruvate kinase, which catalyses the final and

rate limiting step of glycolysis.[14,15] Mammals have four, differentially expressed, pyruvate kinase isoforms, encoded by two genes- PKL (pyruvate kinase liver isoform) and PKM (pyruvate kinase muscle isoform), located on chromosome 1 and 15, respectively.[16] Christofk *et al.*[14] found that tumour cell lines preferentially express isoform 2 of the PKM transcript (PKM2), which drives aerobic glycolysis, supporting anabolic metabolism. This PKM2 isoform is transcribed from the PKM gene, which also encodes for isoform 1 (PKM1). PKM1 is a highly active enzyme, supporting glucose metabolism by OXPHOS.[14] The PKM gene is alternatively spliced to generate the desired enzyme transcript.

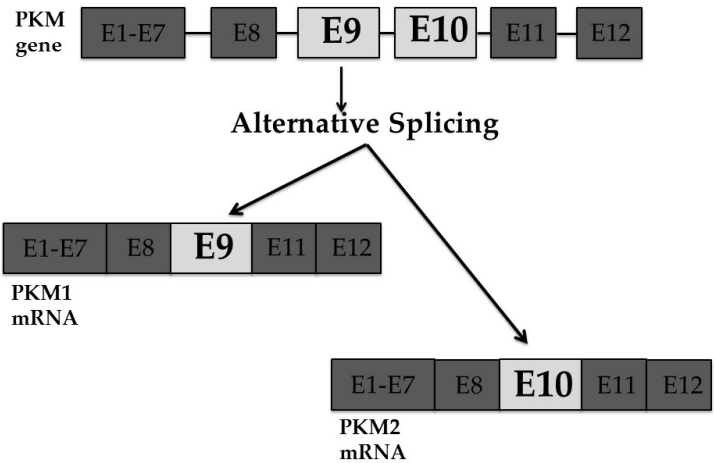


Figure 2: The PKM gene undergoes alternative splicing. The PKM gene is located on chromosome 15. It encodes for two, alternatively spliced isoforms of the pyruvate kinase enzyme, PKM1 and PKM2. These isoforms are differentially expressed and have distinct metabolic roles.
This figure has been adapted with permission from Chaneton et al. (2012).[17]

Cancer cells have developed mechanisms to preferentially splice the PKM2 isoform rather than PKM1. Members of the heterogenous nuclear ribonucleoprotein family flank to exon 9, preventing its inclusion and ensuring transcription of PKM2.[18]

PKM2 exists as both a dimer and a tetramer which have different metabolic roles. Tetrameric PKM2 favours ATP production through the Tricarboxylic Acid Cycle (TCA) as it has a high affinity for its substrate, PEP. Dimeric PKM2 however, is less efficient enzymatically and plays a critical role in aerobic glycolysis and lactate production. Because of its low substrate affinity, the PKM2 dimer results in the accumulation of glycolytic intermediates which are channelled into anabolic pathways, thus rewiring tumour cell metabolism towards biosynthesis, necessary for tumour growth.[19] The PKM2 dimer also has the added ability of

nuclear translocation where it can influence gene expression. Reglation of this dimer:tetramer ratio is important for tumourigenesis, as it is the less active PKM2 dimer which enhances xenograft tumour formation.[20]

PKM2 expression determines the metabolic phenotype of cancer cells, by promoting lactic acid production through aerobic glycolysis, which also influences immune cells in the TME, inducing a pro-tumourigenic macrophage phenotype.[10] Therefore, tumour cell metabolism not only induces tumour growth and development through malignant cell proliferation, but also through macrophage re-education in the TME. This process will now be discussed in detail.

Tumour Cell Metabolism and its Role in Macrophage Re-education and Subsequent Tumour Growth

TAMs are found in high quantities at the periphery of solid tumours and possess an AAM-like phenotype with tumour-promoting functions.[21] TAMs are a plastic group of cells, capable of re-education and functional polarisation with respect to their environmental state. In the TME, the macrophage phenotype associated with cancer initiation is similar to a CAM, displaying pro-inflammatory, anti-tumour responses.[22] However, as the tumour progresses toward malignancy, this phenotype changes to mimic an anti-inflammatory, AAM phenotype, promoting tumour growth.[23] Lactic acid, a metabolic by-product of aerobic glycolysis, has been identified as a key signalling molecule driving macrophage re-education in the TME. Colegio *et al*. showed that tumour-derived lactic acid promotes HIF-1α stabilisation and translocation in surrounding macrophages, resulting in an increased experssion of VEGF and arg-1 which support tumour metastasis and growth, through blood vessel formation and polyamine synthesis (involved in cell growth), respectively.[10] The mechanisms of HIF-1α stabilisation were not discussed, but it is assumed that stabilisation occurs *via* the inhibition of prolyl-hydroxylases (PHDs). Under resting and normoxic conditions, HIF-1α is hydroxylated at conserved proline residues by PHDs. Hydroxylation tags the HIF-1α sububit for ubiquitnation by the von Hippel-Lindau E3 ubiquitin ligase, resulting in rapid proteosomal degradation. PHD inhibition thus results in HIF-1α stabilisation and subsequent nuclear translocation, increasing transcription of desired genes containing hypoxia response elements (HREs), in this case, VEGF and arg-1.[24]

Other tumour-derived molecules, such as adenosine, are known to stabilise HIF-1α. However it was revealed that these signals did not promote VEGF expression, so had no role to play in promoting this pro-tumourigenic TAM phenotype.[10]

In tumour cells, HIF-1α is critical to establish Warburg metabolism, (the metabolic switch from OXPHOS to aerobic glycolysis) as it increases expression of glycolytic machinery in both hypoxic and normoxic environments.[25] Warburg metabolism was primarily thought to be an adaptive mechanism used by tumour cells in response to

hypoxic conditions, due to a decreased blood supply seen in solid tumours. It is now understood that tumour cells also experience this metabolic switch under normoxic conditions. As stated, aerobic glycolysis allows for rapid cellular proliferation by providing metabolic intermediates necessary for biosynthesis.[13] The major metabolic by-product of aerobic glycolysis, lactic acid, is involved in macrophage re-education to promote a pro-tumourigenic phenotype.[10] This suggests that tumour metabolism not only encourages tumour cell proliferation, but too supports solid tumour growth and metastasis by promoting a pro-tumourigenic macrophage phenotype. Re-education of a macrophage in the TME to induce this phenotype, simulates the formation of new blood vessels and polyamine synthesis, resulting from the HIF-1α mediated transcription of VEGF and arg-1 in the TAM. HIF-1α is therefore a major regulator of solid tumour growth, controlling both tumour cell proliferation and macrophage re-education in the TME.

The Complexities of HIF-1α

The metabolic profile of pro-inflammatory CAMs was first investigated by Hard in 1970 were he observed an increased glycolytic flux and a decrease in cellular oxygen consumption.[26] This was supported nearly twenty years later by Newsholme who showed that CAMs displayed an increased expression of glycolytic enzymes, indicating an increase in glycolytic activity.[27] However, it is only recently that researchers have began to appreciate the impact of cellular metabolism on immune function.

It is now established that CAMs utilise aerobic glycolysis as a source of ATP rather than OXPHOS, a phenomenon similar to the Warburg effect, described in tumour cells. This metabolic switch results in the accumulation of TCA intermediates, acting as signals to influence macrophage effector function.[28] In 2013, Tannahill *et al*.[29] investigated the impact of this activation-induced metabolic switch on macrophage effector function. They observed an increase in the TCA intermediate succinate and further revealed that LPS-induced succinate stabilises HIF-1α, under normoxic conditions, with IL-1β being an important transcriptional target. The principal source of succinate in LPS-activated macrophages is through glutamine-dependent anerplerosis. Succinate leads to HIF-1α stabilisation and activation by impairing PHD activity. This effect was inhibited by 2-deoxyglucose, an inhibitor of the glycolytic pathway, suggesting that the metabolic switch to glycolysis is necessary to promote a HIF-1α-induced pro-inflammatory response in CAMs.[29] The mechanisms of this HIF-1α driven pro-inflammatory macrophage phenotype were further clarified in 2015 by Palsson-McDermott *et al*.[30] They found, that LPS increases the expression of the key glycolytic regulator, the M2 isoform of pyruvate kinase-PKM2, which translocates to the nucleus as a dimer alongside HIF-1α, resulting in IL-1β transcription. HIF-1α-induced IL-1β transcription is dependent upon PKM2 binding and translocation, as forcing PKM2 subunits into tetramers (ultimately preventing nuclear translocation), using small molecule activators DASA-58 and

TEPP-46[31,32] attenuated the LPS-induced pro-inflammatory phenotype, causing decreased IL-1β transcription and boosting expression of the anti-inflammatory cytokine, IL-10.[30]

HIF-1α is a complex transcription factor, able to induce both pro-inflammatory and pro-tumourigenic macrophage phenotypes. It is necessary to understand the differential mechanisms driving each HIF-1α-induced phenotype, to exploit these processes clinically. Clarifying the differences between both HIF-1α-driven phenotypes could lead to potential therapeutic targets to prevent macrophage re-education in the TME, increasing the percentage of anti-tumour macrophages, subsequently leading to tumour death.

The Role of PKM2 in Macrophage Re-education?

Extensive evaluation of the literature suggests that the role of PKM2 in promoting the TAM phenotype has not been considered. However, Palsson-McDermott *et al*.[30] noticed that activation of PKM2 by promoting tetramerization, attenuated an LPS-induced pro-inflammatory phenotype while promoting anti-inflammatory macrophage characteristics such as expression of IL-10.[30] Fig. 3 highlights the different metabolic profiles of a resting macrophage, a CAM and a CAM following PKM2 tetramerization.

Future research is necessary to assess the role of PKM2 in macrophage re-education by tumours, as it is possible that tumour secretions may tamper with PKM2 dimer:tetramer ratios in order to induce a pro-tumourigenic phenotype. Tumour-derived lactic acid possibly prevents an anti-tumour phenotype from thriving in the TME by driving PKM2 tetramerization in the TAM, thus preventing HIF-1α-induced IL-1β transcription and a pro-inflammatory response. Another possibility is that tumour secretions down-regulate PKM2 and promote PKM1 expression, in contrast to LPS which results in increased PKM2 expression in CAMs. As stated, the M1 isoform of the PKM gene- PKM1- supports glucose metabolism through OXPHOS, the metabolic profile of anti-inflammatory, AAMs.[14] Upregulation of PKM1 and decreased expression of PKM2 may be the mechanism by which tumour cells enhance an AAM-like phenotype in their surroundings.

It's known that PKM2 tetramerization induces IL-10 expression, however the role of PKM2 in promoting TAM markers, VEGF and Arg-1 is unknown.[30] PKM2 may serve as the key difference between both HIF-1α-induced phenotypes, due to its selectivity in promoting pro-inflammatory gene transcription. Understanding whether tumour-derived signals interfere with PKM2 expression/translocation alongside HIF-1α in a TAM is necessary for TAM-targetted cancer therapy to potentially prevent macrophage re-education and subsequent tumour growth.

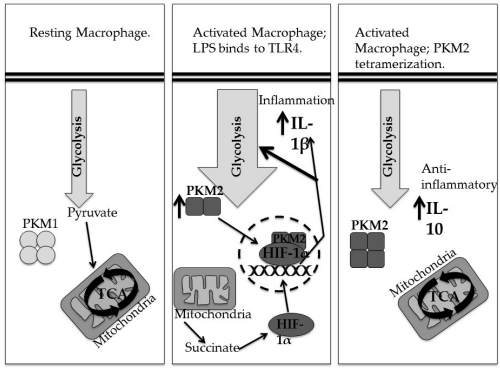


Figure 3: The metabolic profiles of a resting macrophage, a CAM and a macrophage activated following PKM2 tetramerization. A resting macrophage obtains its energy through OXPHOS. Similar to the Warburg effect seen in tumour cells, CAMs undergo a metabolic switch to aerobic glycolysis, resulting in the accumulation of TCA intermediates to aid the pro-inflammatory response. LPS also increases expression of the PKM2 isoform of pyruvate kinase, which in a dimer form, can translocate alongside HIF-1α and enhance IL-1β transcription. Macrophage activation following PKM2 teteramerization using small molecule activators to prevent its nuclear translocation, attenuates the LPS-induced pro-inflammatory response while promoting traits typical of an AAM, highlighting the importance of dimeric PKM2 in driving inflammation.

This diagram has been adapted with permission from Palsson-McDermott et al. (2015).[30]

TAMs as a Potential Target for Cancer Therapy

Cancer immunotherapy encompasses the idea of arming the host's immune system to stimulate an anti-tumour response. Early immunotherapy to increase activation of tumour-infiltrating lymphocytes has been shown to increase patient survival rates in metastatic melanoma.[33] As macrophages constitute the first line of defense, they represent a clever tool with the potential to be manipulated to destroy tumour cells early in the disease course. The objective of macrophage-targetted immunotherapy is to prevent macrophage re-education in the TME, thus enhancing a pro-inflammatory, anti-tumour response.[21] Modulating the macrophage response in the TME to promote a phenotype similar to that of CAMs would result in direct killing of malignant cells, but too, would stimulate a Th1-type cytotoxic T cell response to boost tumour death.[34]

A recent publication in the Journal of Clinical Investigation describes an antibody based therapy to prevent macrophage inhibition by tumour cells in small cell lung cancer (SCLC) cell lines. The Sage lab showed that CD47-blocking antibodies, prevent its engagement with SIRPα, an inhibitory receptor on macrophage cell surfaces. Blocking this interaction increased the phagocytic activity of macrophages in cell culture. What is most exciting is that in a mouse model, administration of CD47-specific antibodies, inhibited the growth of SCLC tumours, highlighting this

CD47-SIRPα interaction as a potential immunotherapeutical target.[35]

I believe that future research must be conducted in order to determine the molecular mechanisms driving macrophage re-education by tumours, which may lead to a promising therapeutic target in cancer immunotherapy.

Conclusion

TAMS are linked with a poor clinical prognosis, perhaps due to their role in tumour growth and progression. In order to prevent tumour driven macrophage re-education in the TME, it is essential to understand the mechanisms promoting this TAM phenotype. HIF-1α is a cricial player in this process, however its potential binding partners have not been established. As PKM2 is essential to support HIF-1α-induced IL-1β expression, the role of PKM2 in macrophage re-education by tumours must be assessed. Understanding the different mechanisms behind each HIF-1α-driven phenotype could uncover a valuable approach to prevent macrophage re-education in the TME, promoting an anti-tumour macrophage phenotype, leading to tumour destruction.

Acknowledgements

This review has been written in preparation for a final year research project for the module PG4902. I would like to thank Dr. Eric Downer, Dr. Frederick Sheedy and Dr. Zbigniew Zaslona for their continued support throughout the year.

References

1. Sica A, Mantovani A. Macrophage plasticity and polarization: in vivo veritas. *J Clin Invest* [Internet]. 2012 Mar;122(3):787–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22378047

2. Murray PJ, Wynn TA. Protective and pathogenic functions of macrophage subsets. *Nat Rev Immunol* [Internet]. 2011 Oct 14;11(11):723–37. Available from: http://www.nature.com/doifinder/10.1038/nri3073

3. Kelly B, O'Neill LAJ. Metabolic reprogramming in macrophages and dendritic cells in innate immunity. *Cell Res* [Internet]. 2015 Jul;25(7):771–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26045163

4. Mills EL, O'Neill LA. Reprogramming mitochondrial metabolism in macrophages as an anti-inflammatory signal. *Eur J Immunol* [Internet]. 2016 Jan;46(1):13–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26643360

5. Clària J, González-Périz A, López-Vicario C, Rius B, Titos E. New insights into the role of macrophages in adipose tissue inflammation and fatty liver disease: modulation by endogenous omega-3 fatty acid-derived lipid mediators. *Front Immunol* [Internet]. 2011;2:49. Available from: http://journal.frontiersin.org/article/10.3389/fimmu.2011.00049/abstract

6. Pollard JW. Trophic macrophages in development and disease. *Nat Rev Immunol* [Internet]. 2009 Apr;9(4):259–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19282852
7. Balkwill FR, Capasso M, Hagemann T. The tumour microenvironment at a glance. *J Cell Sci* [Internet]. 2012;125(23):5591–6. Available from: http://jcs.biologists.org/cgi/doi/10.1242/jcs.116392

8. Quatromoni JG, Eruslanov E. Tumour-associated macrophages: function, phenotype, and link to prognosis in human lung cancer. *Am J Transl Res* [Internet]. 2012;4(4):376–89. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23145206

9. Quail DF, Joyce JA. Microenvironmental Regulation of Tumour Progression and Metastasis. *Nat Med*. 2013;19(11):1423–37.

10. Colegio OR, Chu N-Q, Szabo AL, Chu T, Rhebergen AM, Jairam V, *et al.* Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature* [Internet]. 2014 Jul 13;513(7519):559–63. Available from: http://www.nature.com/doifinder/10.1038/nature13490

11. Farkona S, Diamandis EP, Blasutig IM. Cancer immunotherapy: the beginning of the end of cancer? *BMC Med* [Internet]. 2016;14:73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27151159%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4858828

12. Warburg O. On the Origin of Cancer Cells. *Source Sci New Ser* [Internet]. 1956;123(3191):309–14. Available from: http://www.jstor.org

13. Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol* [Internet]. 2015 Apr;17(4):351–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25774832
14. Christofk HR, Vander Heiden MG, Harris MH, Ramanathan A, Gerszten RE, Wei R, *et al*. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* [Internet]. 2008 Mar 13;452(7184):230–3. Available from: http://www.nature.com/doifinder/10.1038/nature06734

15. Israelsen WJ, Vander Heiden MG. Pyruvate kinase: Function, regulation and role in cancer. *Semin Cell Dev Biol* [Internet]. 2015 Jul;43:43–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26277545

16. Jurica MS, Mesecar A, Heath PJ, Shi W, Nowak T, Stoddard BL. The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure*. 1998;6(2):195–210.

17. Chaneton B, Gottlieb E, Mazurek S, Harada K, al. et, Noguchi T, *et al*. Rocking cell metabolism: revised functions of the key glycolytic regulator PKM2 in cancer. *Trends Biochem Sci* [Internet]. 2012 Aug;37(8):309–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22626471

18. Chen M, Zhang J, Manley JL. Turning on a Fuel Switch of Cancer: hnRNP Proteins Regulate Alternative Splicing of Pyruvate Kinase mRNA. *Cancer Res.* 2010;70(22).

19. Dayton TL, Jacks T, Vander Heiden MG. PKM2, cancer metabolism, and the road ahead. *EMBO Rep* [Internet]. 2016 Dec;17(12):1721–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27856534

20. Wong N, De Melo J, Tang D. PKM2, a Central Point of Regulation in Cancer Metabolism. *Int J Cell Biol* [Internet]. 2013;2013:242513. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23476652

21. Weagel E, Smith C, Liu PG, Robison R, Neill KO'. Macrophage Polarization and Its Role in Cancer. *J Clin Cell Immunol.* 2015;6(4).

22. Gordon S. Alternative activation of macrophages. *Nat Rev Immunol* [Internet]. 2003 Jan;3(1):23–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12511873

23. Qian B-Z, Pollard JW, Abraham D, Zins K, Sioud M, Lucas T, *et al*. Macrophage diversity enhances tumour progression and metastasis. *Cell* [Internet]. 2010 Apr 2;141(1):39–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20371344

24. Corcoran SE, O'Neill LAJ. HIF1α and metabolic reprogramming in inflammation. *J Clin Invest* [Internet]. 2016 Aug 29;126(10):3699–707. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27571407

25. Robey IF, Lien AD, Welsh SJ, Baggett BK, Gillies RJ. Hypoxia-inducible factor-1alpha and the glycolytic phenotype in tumours. *Neoplasia*. 2005;7(4):324–30.

26. Hard GC. Some biochemical aspects of the immune macrophage. *Br J Exp Pathol* [Internet]. 1970 Feb;51(1):97–105. Available from: http://www.ncbi.nlm.nih.gov/pubmed/5434449

27. Newsholme P, Gordon S, Newsholme EA. Rates of utilization and fates of glucose, glutamine, pyruvate, fatty acids and ketone bodies by mouse macrophages. *Biochem J* [Internet]. 1987 Mar 15;242(3):631–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3593269

28. Jha AK, Huang SC-C, Sergushichev A, Lampropoulou V, Ivanova Y, Loginicheva E, *et al*. Network Integration of Parallel Metabolic and Transcriptional Data Reveals Metabolic Modules that Regulate Macrophage Polarization. *Immunity* [Internet]. 2015 Mar;42(3):419–30. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1074761315000801

29. Tannahill GM, Curtis AM, Adamik J, Palsson-McDermott EM, McGettrick AF, Goel G, *et al*. Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* [Internet]. 2013 Mar 24;496(7444):238–42. Available from: http://www.nature.com/doifinder/10.1038/nature11986

30. Palsson-McDermott EM, Curtis AM, Goel G, Lauterbach MAR, Sheedy FJ, Gleeson LE, *et al*. Pyruvate Kinase M2 Regulates Hif-1α Activity and IL-1β Induction and Is a Critical Determinant of the Warburg Effect in LPS-Activated Macrophages. *Cell Metab* [Internet]. 2015 Jan 6;21(1):65–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25565206

31. Boxer MB, Jiang J, Vander Heiden MG, Shen M, Skoumbourdis AP, Southall N, *et al*. Evaluation of Substituted N, N'-Diarylsulfonamides as Activators of the Tumour Cell Specific M2 Isoform of Pyruvate Kinase. *J Med Chem* [Internet]. 2010 Feb 11;53(3):1048–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20017496

32. Jiang J, Boxer MB, Vander Heiden MG, Shen M, Skoumbourdis AP, Southall N, *et al*. Evaluation of thieno[3,2-b]pyrrole[3,2-d] pyridazinones as activators of the tumour cell specific M2 isoform of pyruvate kinase. *Bioorg Med Chem Lett* [Internet]. 2010 Jun 1;20(11):3387–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20451379

33. Morton DL, Foshag LJ, Hoon DS, Nizze JA, Famatiga E, Wanek LA, *et al*. Prolongation of survival in metastatic melanoma after active specific immunotherapy with a new polyvalent melanoma vaccine. *Ann Surg* [Internet]. 1992 Oct;216(4):463–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1417196

34. Mills CD, Lenz LL, Harris RA. A Breakthrough: Macrophage-Directed Cancer Immunotherapy. *Cancer Res* [Internet]. 2016 Feb 1;76(3):513–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26772756

35. Weiskopf K, Jahchan NS, Schnorr PJ, Cristea S, Ring AM, Maute RL, *et al*. CD47-blocking immunotherapies stimulate macrophage-mediated destruction of small-cell lung cancer. *J Clin Invest* [Internet]. 2016 Jun 13;126(7):2610–20. Available from: https://www.jci.org/articles/view/81603

TS
SR

Viral Immune Evasion: The STING in the Tale

Dearbhla Murphy
Junior Sophister
Immunology

The cGAS-STING pathway is an intracellular sensing pathway that has evolved to detect and respond to the presence of viral DNA within the cytosol. In response to this, viruses have evolved to evade certain immune responses, like those solicited by STING. They do this through a number of mechanisms, such as the suppression of the cGAS-STING pathway or the manipulation of cellular components such as the exonuclease Trex1. This demonstrates how viruses have coevolved with their host cells to carry on using their cellular machinery to produce infectious progeny. Many advances have been made in this area, yet it is still not fully understood. Gaining an insight into how viruses subvert immune responses is crucial to understanding and subsequently treating viral diseases.

Introduction

When a cell is invaded by a pathogen, such as a virus, the innate immune system is triggered and rapidly becomes activated in order to deal with the invading non-self. Innate immunity is an important mechanism by which the body keeps infection and illness at bay. The recognition of foreign or "non-self" materials, such as microbial nucleic acids, is one of the major mechanisms employed by the immune system to detect pathogens. Cytosolic GMP-AMP synthase (cGAS) is an important sensor molecule for cytosolic DNA, particularly double stranded viral DNA.[1] Upon DNA binding, cGAS undergoes a conformational change, which leads to the activation of the enzyme and the synthesis of a second messenger, 2'3'-cyclic GMP–AMP (cGAMP) from ATP and GTP, which is a ligand for the stimulator of interferon genes (STING).[2] This causes STING to dimerise and induces the production of type 1 interferons (IFN) through the activation of a number of transcription factors, TBK1, IRF3, IRF7 and NFkB, as seen in the diagram below. Type 1 IFNs are essential for fighting off intracellular pathogens, especially viruses.[3]

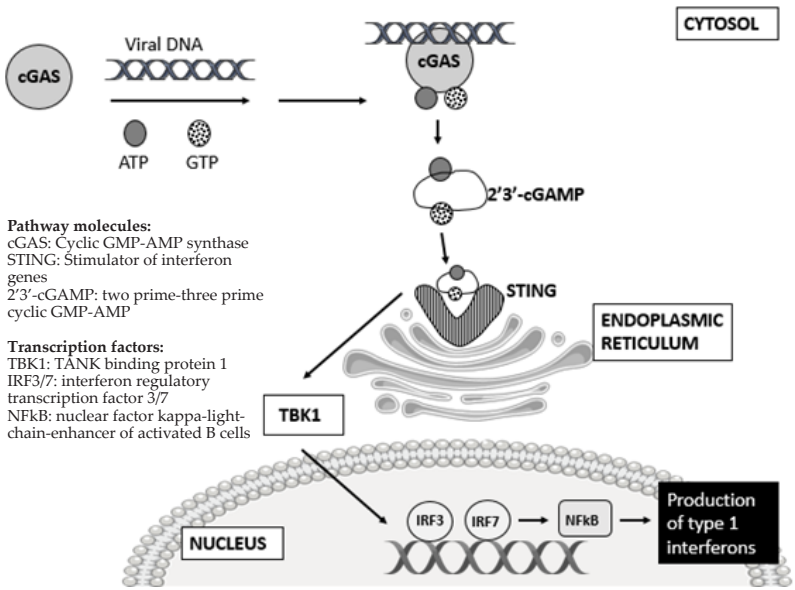


Figure 1: The cGAS-cGAMP-STING pathway. The messenger molecule Cyclic GMP-AMP binds to and subsequently activates STING causing the induction of type I interferons. cGAMP binds the STING dimer like a cap and causes it to phosphorylate TBK1, activating IRF7 and IRF3. IRF3 in the nucleus of the cell where it and NFkB induce type I interferons against the host invading virus.[29]
Adapted from "The cGAS-cGAMP-STING pathway of cytosolic DNA sensing and signalling"

The Stimulator of Interferon Genes (STING)

STING has evolved to be an effective viral sensor, recognising DNA in the cytosol of the cell and responding by inducing an anti-viral response in surrounding cells through the production of type I IFNs. The secretion of type I IFNs helps prevent the spread of viral infections by provoking an immune response against the invading virus through the production of cytokines.[4] Cytokines are substances secreted by a variety of cell types, such as immune cells and endothelial cells, and have an effect on other cells, such as type 1 IFNs, interleukins or even growth factors. In response to this certain viruses have evolved to evade the immune system. The co-evolution of viruses with their host cells has led to the emergence of viral pathogens which can evade detection by the pathogen recognition receptors (PRRs), the sensors of the innate immune system. Similarly, viruses have also evolved to suppress immune responses so that they may replicate their genome within the host cell and produce progeny, continuing their genetic lineage.[4] This review details the mechanisms by which viruses have co-evolved with the viral sensing molecule STING to evade immune detection and retaliation.

Viruses and STING

Viruses are obligate parasites, meaning they require the use of host cell machinery as they do not possess their own, which is needed to replicate their genome.[5] Viruses enter their host and infect their host cells in order to replicate the genes needed for them to produce daughter viral cells. Some viruses require the use of host cell polymerases to replicate their genome and others can encode their own replication factors. The replication of any viral genome is highly dependent on a cellular state which allows for DNA replication and, thus their use of cell components relies heavily on the cell cycle.[6] For example, a virus may induce the cell to forcefully undergo cell division, which may lead to transformation of the cell and, ultimately, lead to cancer. In response, mammalian cells have evolved elaborate intracellular and extracellular mechanisms to detect and inhibit viral replication. This has subsequently led to the emergence of viruses that are able to manipulate and undermine most host immune responses, allowing them to replicate and spread to new tissues and hosts undisturbed.[7] Examples of such viral strategies include the modification of viral nucleic acids, the degradation or cleavage of PRRs and their adaptor proteins, interference with specific post-translational modifications of PRRs and their adaptor proteins (such as STING) and the re-localisation of PRRs within the host cell.[8]

The cGAS-STING pathway detects viruses by the presence of dsDNA in the cytosol of the cell which has spilled out from the nucleus during viral replication. cGAS binds this DNA and synthesises cGAMP, which then binds STING, leading to its activation and the initiation of an anti-viral response.[9] cGAMP can also travel to neighbouring cells through channels out of the cell known as gap junctions enabling cGAS to induce an anti-viral response in the surrounding area and cells, thus limiting the spread of the virus within the body.[10] However, viruses have evolved to evade and suppress important immune molecules like STING and the subsequent production of type I IFNs by targeting important transcription factors in the cGAS-STING pathway, such as TBK1, IRF3, IRF7 and NF-kB. These transcription factors stimulate the expression of the genes responsible for type I IFN production. Viruses have also evolved to block the function of specific antiviral effector proteins and type 1 interferons (IFNα/β) receptor signalling, which impedes the initiation of an immune response against them.[11]

Regulation of Post Transcriptional Modifications

Viral manipulation of STING and the post translational modification of STING are a means viruses have developed to suppress or evade the immune response. An example of a virus which has evolved to evade detection by STING in this manner is the Hepatitis B virus (HBV).[12] HBV is a major cause of chronic hepatitis and infection with the virus increases the risk of developing hepatocellular carcinoma and cirrhosis of the liver.[8] The addition of a ubiquitin molecule to a substrate

protein, such as STING, is called ubiquitination. When cGAMP binds STING, it causes it to dimerize and undergo Lys63-linked ubiquitination, which is crucial for the activation of STING.[8] Recent studies have shown that the reverse transcriptase and ribonuclease H domains of the HBV polymerase are able to bind to STING and subsequently block its Lys63-linked ubiquitination. This results in the inhibition of the production of IFN-β.[12] Without IFN-β, no anti-viral response can be initiated and the HBV can continue its replication cycle.

A recent study on viral immune evasion has shown that DNA tumour viruses such as the human adenovirus 5 (hAd5) and the human papillomavirus 18 (HPV18) inhibit the cGAS–STING pathway through the use of their viral oncoproteins E7 and E1A, respectively.[14] Viral oncoproteins are protein molecules coded for by a viral oncogene which has been integrated into the genome of a eukaryotic cell.[15] These oncoproteins are involved in the regulation and synthesis of proteins linked to tumour cell growth.[16] E7 and E1A are thought to specifically bind to STING but not to cGAS. However, the work in question did not explore whether the ubiquitination or the dimerisation of STING, both of which are induced by the binding of 2′3′-cGAMP, were also affected by this oncoprotein binding.[8,14,15] LXCXE, a motif in the viral oncoproteins, has been shown to be essential to the viral inhibition of the retinoblastoma tumour suppressor pathway, and is also the motif responsible for interaction with STING.[8,14,15] This finding suggests that the inactivation of tumour suppressor function of STING and the antagonism of innate immune signalling share common mechanistic features.[8] As many DNA viruses which invade humans encode oncoproteins with LXCXE motifs, the study proposed that these virally encoded oncoproteins could be representative of a broad class of STING antagonists.[8] If the pathway is inhibited, viruses are able to replicate without cytosolic detection.

The papain-like proteases (PLPs) from human coronavirus NL63, severe acute respiratory system coronavirus (SARS-CoV) and porcine epidemic diarrhoea virus (PEDV), all of which are RNA viruses, are also capable of suppressing the immune response of STING. PLPs are enzymes which break down proteins and peptides and have a papain-like fold.[17] These RNA viruses have been shown to associate with STING by blocking its dimerization and subsequently, Lys63-linked ubiquitination.[18,19,20] This results in the suppression of the activation of IRF3, an important transcription factor for the production of IFNs.

Kaposi's sarcoma-associated herpesvirus (KSHV) is the causative agent of a number of lymphoproliferative diseases in humans, such as Kaposi sarcoma. KSHV is also an encoder of viral interferon regulatory factor 1 (vIRF1). This regulatory factor is able to interact with STING, blocking its binding of the transcription factor TBK1 and thus, TBK1-mediated phosphorylation of STING. This means that STING cannot be activated and produce downstream effects.[8,21] Upon screening KSHV proteins for their ability to inhibit the STING pathway, six viral proteins that block IFN-β activation through this pathway were identified.[21]

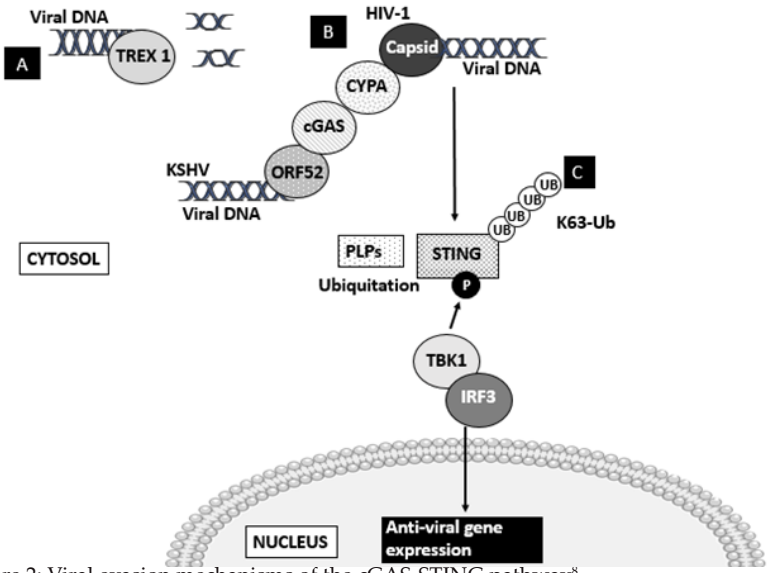


Figure 2: Viral evasion mechanisms of the cGAS-STING pathway[8]

(A) Viral DNA blocking TREX1, meaning that self ssDNA cannot be broken down.
(B) HIV-1 and KSHV also bind cGAS, meaning it cannot activate STING.
(C) Blocking of the dimerization and the ubiquitination of STING by PLPs. This blocks the phosphorylation of TBK1, an important TF in the pathway[8]

Adapted from "Viral evasion of intracellular DNA and RNA sensing"

Viral Sequestration Mechanisms

Sequestration is the accumulation of a compound or tissue, such as viral genomes. It is another means by which viruses have evolved to evade immune detection. The human immunodeficiency virus (HIV) is an example of such a virus. The virus particles of the HIV-1 carry a single stranded RNA genome. This is reverse-transcribed into DNA in the infected host cell in order to prepare the virus for chromosomal integration with its host cell using the viral enzyme integrase. As only a few of the many viral DNA strands integrate with the host DNA, many viral DNA strands end up in the cytosol where they might be detected by cytosolic sensors like cGAS. As such, the virus makes use of the host repair exonuclease TREX1, an encoder of a DNA exonuclease needed to degrade self ssDNA and dsDNA substrates. TREX1 binds to and subsequently degrades HIV-1 DNA and thus avoids an immune reaction.[22] Undetected viruses can go on replicating their genome and infect surrounding cells.

Relocation of DNA Sensors

Another mechanism of evasion and suppression employed by viruses is the cleavage or degradation of DNA sensors, like cGAS, and also the STING molecule itself. Dengue virus (DENV) is a ssRNA virus and the cause of dengue fever. Dengue fever is a disease transmitted by mosquitoes and causes sudden fever and acute pain in the joints. The protease complex N52B–N53 of the dengue virus (DENV) binds and cleaves human STING resulting in its inactivation and in reduced induction of IFNs.[23] STING is deemed a species-specific restriction factor of DENV replication, whereas the NS2B-Ns3 fails to cleave STING in mice. This is because the murine STING does not have the specific cleavage site for DENV that human STING does. This shows not only how viruses coevolved with STING but they have also done so in a species specific manner. This, along with other previous studies, highlights the convergence of DENV proteases on adaptors of both RNA and DNA sensors.[8,23]

Finally, the relocalisation of DNA sensors is also utilised by viruses when evading an immune response by STING. A recent study showed that the enveloping membrane of the protein ORF52 of KSHV was a potent inhibitor of cGAS signalling. It demonstrated that the protein ORF52 of KSHV binds to DNA and also to cGAS, blocking the enzymatic activity of cGAS, which facilitates the evasion of the host IFN response.[24] This was proven by individually assaying each KSHV open-reading frame for its ability to impair activation of an IFN reporter gene which was driven by cGAS activity.[24] Out of all the KSHV ORF proteins which reduced IFN reporter stimulation, only ORF52 seemed to display both cytoplasmic location and DNA-binding activity.[24,25] Further to this discovery, the authors also found that ORF52 homologues from other species of gammaherpesvirus such as the Epstein Barr virus (EBV) and Murine Gammaherpesvirus 68 (MHV68), also antagonised cGAS activity.[8,24] This finding supports a conserved immune evasion function for the ORF52 protein. Alongside viRF1, another protein from KSHV which binds and antagonises STING, it can be concluded that gamma herpes viruses target both the sensor and adaptor molecule of the cGAS-STING pathway for effective immune evasion.[24]

Discussion and Conclusion

Huge advances in understanding the mechanisms of viral evasion of the immune response have been made in the past decade, especially with regards to the STING molecule. However, there are still notable gaps in our knowledge. For example, both RNA and DNA viruses such as DENV and HBV antagonize intracellular viral RNA and DNA receptors, indicating crosstalk between these two major sensing pathways. This interaction remains quite poorly defined.[8] More research needs to be carried on the sensing capabilities of STING especially in regards to the sensing of viral RNA. Some new studies have concluded that cGAS-STING pathway is

capable of recognising RNA viruses[26,27] however more research into this area is needed.

Considering all of the recent findings on the cGAS-STING pathway and DNA sensing, it is evident that this is clearly both an exciting and rapidly evolving area of immunology. However, in terms of viral immune evasion and the evolution of the STING pathway, many questions still remain. KSHV, like all herpesviruses, deposits its DNA genome into the nucleus of the host cell where it may be replicated by viral DNA polymerase to produce infectious progeny. Therefore, when we look at innate viral sensing, it's difficult to imagine how it's sensed by cGAS, a cytoplasmic DNA sensor. It has been proposed that the disruption of the viral capsule within the host cells' cytoplasm may explain this phenomenon, but this remains to be proven.[25]

Whilst it can be seen that the mechanism of the cGAS-STING pathway is effective at suppressing viral invasion and triggering an immune response, there are still many limitations to its ability to detect all viruses. cGAS has been observed to be able to interact with IFI16, another DNA sensor, promoting an IFI16-dependent response to viruses like the herpesvirus.[25] From this observation we can hypothesise that there is likely additional components to this pathway and crosstalk between pathways which require further investigation and characterisation.

Even though there are still many gaps in our knowledge with respect to cGAS-STING function within an infected cell, the numerous and varying immune evasion strategies employed by not only DNA viruses, but also RNA viruses, against it suggests the importance of this pathway in viral pathogen sensing. So far only STING antagonists from positive-sense RNA viruses and DNA viruses have been identified, but not negative-sense RNA viruses. This starts a conversation about the diversity of STING and what it can sense.[27] During immune evasion, some viruses, both RNA and DNA, use similar mechanisms to inhibit STING function and also STING-TBK1 interactions, while other viruses make use of more unique mechanisms involving STING cleavage and degradation.[27] This shows the diverse way in which viruses have evolved with STING to evade detection by cGAS.

The next step in the evolution of both cGAS and STING is finding a way to undermine these methods of immune evasion employed by viruses. The DENV viral evasion causes the host to contract dengue fever which is a very dangerous disease. Similarly, many viruses can cause cancers through the manipulation of the host's cell cycle causing rapid proliferation or by viral oncoproteins as discussed previously. Combatting the ways in which these viruses undermine the immune system could be important in combatting virally induced cancers such as Kaposi sarcoma caused by KSHV. Likewise, these mechanisms of immune evasion and manipulation could be used when medically combatting viral infection through the use of drugs and vaccination. It has been estimated that viruses are causative agents in approximately 12% of all cases of human cancers. Most of these cancers can be attributed to infections by HPV, HBV, EBV, and KSHV, which are all viruses

capable of evading immune responses.[28] Prophylactic vaccines against other pathogenic viruses have been seen to have a wonderful record as public health interventions in terms of their safety, ability to reach economically disadvantaged populations and effectiveness in combatting viral disease. Hopefully these considerations will inspire efforts to develop and implement vaccines against oncoviruses, especially in combatting their ability to subvert the immune system.[28] The cGAS-STING pathway and the viral evasion of this pathway is a rapidly expanding area of interest in Immunology which could prove very relevant in the future treatment and prevention of viral diseases.

Acknowledgements

This review was adapted from a previous review "Innate Viral Sensing: The STING in the Tale" submitted for the JS Immunology Research Skills module (BI3020). Figures were adapted from referenced papers.

References

1. Xiao TS, Fitzgerald K. The cGAS-STING Pathway for DNA Sensing. *Molecular Cell.* 2013;51(2):135–9.

2. Christensen MH, Paludan SR. Viral evasion of DNA-stimulated innate immune responses. *Cell Mol Immunol.* 2016;(October 2015):1–10.

3. Sun L, Wu J, Du F, Chen X, Chen ZJ. Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway. *Science* (80-). 2013;339(6121).

4. Bowie AG, Unterholzner L. Viral evasion and subversion of pattern-recognition receptor signalling. *Nat Rev Immunol*. 2008;8(12):911–22.

5. Miller S, Krijnse-Locker J. Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol.* 2008;6(5):363–74.

6. Doorbar J. The papillomavirus life cycle. *J Clin Virol.* 2005;32(SUPPL.):7–15.

7. Taylor KE, Mossman KL. Recent advances in understanding viral evasion of type I interferon. *Immunology.* 2013;138(3):190–7.

8. Kai Chan Y, Gack MU. Viral evasion of intracellular DNA and RNA sensing. *Nat Publ Gr.* 2016;14(6):360–73.

9. Ishikawa H, Barber GN. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature*. 2008;455(7213):674–8.

10. Ablasser A, Schmid-Burgk JL, Hemmerling I, Horvath GL, Schmidt T, Latz E, *et al*. Cell intrinsic immunity spreads to bystander cells *via* the intercellular transfer of cGAMP. *Nature*. 2013;503(7477):530–4.

11. Bowie AG, Unterholzner L. Viral evasion and subversion of pattern-recognition receptor signalling. 2008;8(December).

12. Liu Y, Li J, Chen J, Li Y, Wang W, Du X, *et al*. Hepatitis B Virus Polymerase Disrupts K63-Linked Ubiquitination of STING To Block Innate Cytosolic DNA-Sensing Pathways. *J Virol*. 2015;89(4):2287–300.

13. Chen Q, Sun L, Chen ZJ. Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nat Immunol*. 2016;17(10):1142–9.

14. Lau L, Gray EE, Brunette RL, Stetson DB. DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. *Science* (80-). 2015;350(6260):568–71. l.pdf

15. Ledl A, Schmidt D, Müller S. Viral oncoproteins E1A and E7 and cellular LxCxE proteins repress SUMO modification of the retinoblastoma tumor suppressor. *Oncogene*. 2005;24(23):3810–8.

16. Tomaic V. Functional roles of E6 and E7 oncoproteins in HPV-induced malignancies at diverse anatomical sites. *Cancers (Basel)*. 2016;8(10).

17. Baez-Santos YM, St. John SE, Mesecar AD. The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds. *Antiviral Res*. 2015;115:21–38.

18. Chen X, Yang X, Zheng Y, Yang Y, Xing Y, Chen Z. SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex. *Protein Cell*. 2014;5(5):369–81.

19. Sun L, Xing Y, Chen X, Zheng Y, Yang Y, Nichols DB, *et al*. Coronavirus papain-like proteases negatively regulate antiviral innate immune response through disruption of STING-mediated signaling. *PLoS One*. 2012;7(2).

20. Xing Y, Chen J, Tu J, Zhang B, Chen X, Shi H, *et al*. The papain-like protease of porcine epidemic diarrhea virus negatively regulates type I interferon pathway by acting as a viral deubiquitinase. *J Gen Virol*. 2013;94(PART7):1554–67.

21. Ma Z, Jacobs SR, West JA, Stopford C, Zhang Z, Davis Z, *et al*. Modulation of the cGAS-STING DNA sensing pathway by gammaherpesviruses. *Proc Natl Acad Sci U S A*. 2015;112(31):E4306-15.

22. Yan N, Regalado-Magdos A, Stiggelbout B, Lee-Kirsch M, Lieberman J. The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nat Immunol*. 2010;11(September):1005–13.

23. Aguirre S, Maestre AM, Pagni S, Patel JR, Savage T, Gutman D, *et al*. DENV Inhibits Type I IFN Production in Infected Cells by Cleaving Human STING. *Vol. 8, PLoS Pathogens*. 2012.

24. Wu JJ, Li W, Shao Y, Avey D, Fu B, Gillen J, *et al*. Inhibition of cGAS DNA Sensing by a Herpesvirus Virion Protein. *Cell Host Microbe*. 2015;18(3):333–44.

25. Diner BA, Cristea IM. Blowing off Steam: Virus Inhibition of cGAS DNA Sensing. *Cell Host Microbe*. 2015;18(3):270–2.

26. Schoggins JW, MacDuff DA, Imanaka N, Gainey MD, Shrestha B, Eitson JL, *et al*. Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature*. 2014;505(7485):691–5.

27. Ma Z, Damania B. The cGAS-STING Defense Pathway and Its Counteraction by Viruses. *Cell Host Microbe*. 2016;19(2):150–8.

28. Schiller JT, Lowy DR. Vaccines to Prevent Infections by Oncoviruses *. *Annu Rev Microbiol*. 2010;64(1):23–41.

29. Cai X, Chiu YH, Chen ZJ. The cGAS-cGAMP-STING pathway of cytosolic DNA sensing and signaling. *Mol Cell* [Internet]. 2014;54(2):289–96.

TS
SR

Reconsidering the Role of Mitochondrial Reactive Oxygen Species (mtROS) in Aging and Age-related Disease

Aisling Callaghan
Senior Sophister
Genetics

The biological process of aging is associated with the accumulation of deleterious mutations, the gradual deterioration of physiological function and the increased prevalence of degenerative diseases. The past few decades of research have provided evidence to support the association between the accumulation of mutations in mitochondrial DNA (mtDNA) and the aging process in mammals. This research has been driven by the development of animal models, which have been designed to significantly increase the number of mtDNA mutations accumulated over time. Mitochondrial mutations have been well described in terms of decline in mitochondrial structure and function. Much of the recent work in aging research has cast doubt on the traditional mitochondrial free radical theory of aging which proposes that the free radicals produced as a by-product of mitochondrial metabolism damage mtDNA altering mitochondrial function, which, in turn leads to increased production of reactive oxygen species (ROS) and mitochondrial mutations. From this theory it was predicted that there would be a negative correlation between oxidative damage to mtDNA and aging. However, this theory has been heavily criticised in recent times with further evidence suggesting this is an over-simplification of the aging process. There is now a large body of evidence available which suggests that mitohormesis and mitochondrial reactive oxygen species (mtROS) may promote longevity and mitochondrial health.

Introduction

Aging is considered to be a natural process that is universal to almost all living organisms.[1] Traditionally, aging was thought to be a default process that occurs due to the natural 'wear and tear' of an organism over time.[1] However, decades of multidisciplinary research have since proven that aging is not simply due to these stochastic changes but is likely regulated by a myriad of molecular and cellular factors.[2] The development of model organisms, particularly the mouse, has allowed

scientists to manipulate and regulate aging in these organisms through genetic engineering.[10] This suggests that aging is not a passive process as once thought, but is much more complex.[2] The rate of aging differs between organisms but also differs between the cells in the organism itself. This 'selective aging' is what accounts for the large amount of pathologies within the aging human population such as cancers, diabetes, metabolic diseases, hypertension, arthritis, cardiovascular diseases and neurodegenerative diseases.[3] A number of genetic, biochemical and behavioural techniques are being used in the attempt to uncover the molecular and cellular factors underlying this 'selective aging' process in model organisms.[1,2,10]

The mitochondria has played a principle role in understanding the natural aging process.[1] This organelle is essential in producing the energy needed for the cell to function.[4] It does this by generating ATP through the process of oxidative phosphorylation (OxPhos). OxPhos is carried out by respiratory chain (RC) complexes (1-IV) and ATP synthase located in the inner mitochondrial membrane.[5] Mitochondrial diseases occur when mitochondrial function and its ability to produce this much needed energy is compromised.[4] Mitochondria regulate a number of metabolic and signalling pathways which form a complex network that can be re-modelled in response to molecular cues.[2] Mitochondrial biogenesis occurs through the stimulation of sirtuin 1 (SIR1) and PGC-1α/β transcriptional coactivators.[22] AMP-activated kinase (AMPK) also stimulates the activation of PGC-1α/β through phosphorylation, suppressing mitophagy – the selective degradation of the mitochondria by autophagy.[22] Disruption to normal mitochondrial function leads to proteotoxic stress as a result of protein misfolding.[5] This proteotoxic stress activates the mitochondrial unfolded protein response (UPRmt) which promotes proteostasis in the mitochondria and longevity of the organism.[5]

The Mutational Burden of the Mitochondrial Genome

A large amount of evidence exists to support the widely accepted hypothesis that the accumulation of mutations in the mitochondrial genome has a principal role in biological aging.[2] Recent studies using aging mice models have supported the direct association between mtDNA mutations and decline in mitochondrial function.[9] Mutator mice are genetically engineered to accumulate and sustain mtDNA mutations at much higher rates than normal.[9] These mice are created by genetically engineering homozygous knock-in mice expressing a version of an essential mtDNA polymerase γ (PolG). As expected, these mice demonstrate widespread mitochondrial defects and develop multiple age-related disease phenotypes with premature aging and reduced lifespan.[10] What was not expected, however, was that neither increased production of ROS or oxidative damage was observed in the mice.[11] The mitochondria produce the majority of ROS as a by-product of oxidative phosphorylation. As the mitochondria are the primary source and main producer of ROS, it was previously thought that the accumulation of mutations in mtDNA results in respiratory chain deficiencies and defects in

mitochondrial metabolism. This idea is related to a theory known as 'The Free Radical Theory of Aging'.[12] Not only has this theory been heavily criticised, but studies have also cast doubt on the causative link between mtDNA mutations and aging. This doubt emerges from the fact that these homozygous mutator mice experience a much greater extent of mitochondrial mutations than that of naturally aging mice who may be heterozygous for PolG efficiency.[11] It has also been noted in recent studies that various mutations in mtDNA can have different effects on lifespan, with not all mutations associated with a decreased lifespan. For example, mice carrying a mutation in mtDNA helicase Twinkle do not show decreased longevity and premature aging, despite eventually demonstrating late onset mitochondrial dysfunction and myopathies.[13]

A Criticism of the 'Mitochondrial Free Radical Theory of Aging''

The Mitochondrial Free Radical Theory of Aging postulates that the generation and the accumulation of ROS in the body drives the aging process, causing molecular damage to both mtDNA and nuclear DNA in the cell which, in turn, leads to loss of mitochondrial function.[12] As the majority of ROS are produced by the mitochondria as a by-product of the decline in OxPhos efficiency, it was thought that the accumulation of mitochondrial mtDNA mutations may lead to respiratory defects and the increased production of ROS.[2] This could then lead to further damage of mtDNA which results in the progressive decline of cellular integrity and aging observed across multiple tissues. Much doubt was cast over the theory after mice that were genetically engineered to exhibit a 2,500-fold higher mtDNA mutational frequency than normal mice were not observed to have increased production of ROS or significantly more oxidative damage.[11] It has been noted in previous studies that animals that have undergone caloric restriction do exhibit increased lifespans and produce fewer mitochondrial ROS (mtROS). These results seem to contradict each other; however, dietary restriction modifies processes other than ROS production including insulin signalling, a pathway known to be implicated in longevity, which may be involved in the increased lifespan observed in these mice.[14] In order to truly prove that ROS is a main driver of aging, all other physiological factors affecting longevity must remain unaltered which is an extremely difficult task to achieve.[1]

Recent work surrounding the biology of the naked mole rat has shown further data that contradicts the mitochondrial free radical theory of aging.[15] The naked mole rat has an unusually long lifespan of up to 25 years compared with only 3-4 years in ordinary rats. Despite their increased longevity, the naked mole rats were shown to have higher levels of oxidative damage than the ordinary control rats. These results appear to be a total contradiction to the free radical theory, which implies that there is a negative correlation between increased ROS production and longevity.[12] It was proposed that this increased longevity in naked mole rats may

be partly due to their low metabolic rate and highly efficient cysteine reduction mechanisms which lead to increased proteostasis in the mitochondria.[15]

This suggests that if cellular damage is caused by oxidative stress in the naked mole rate, this could be at least partly compensated for by these protein repair systems. Thus, their unusually long lifespan may not provide total evidence against the mitochondrial free radical theory of aging. Other studies have shown there are also Nuo-6 mutant strains of *C.elegans* which exhibit greater oxidative stress than control strains, yet also exhibit increased longevity.[16] The researchers then tested the affects of antioxidants on longevity and it was found when the same mutant strains were treated with potent antioxidants they did not show increased longevity but rather showed reduced longevity. It is clear from all of these longevity studies that the process of aging is more complex than previously thought. The evidence implies that there must be underlying mechanisms which contribute to the aging process, and which allow for the observed increase in lifespan, despite relatively high levels of mtROS.

An Overview of Mitohormesis

These observations that contradict the mitochondrial free radical theory of aging can be explained by more recent studies suggesting that mild oxidative stress may actually promote mitochondrial stress responses[17] (Figure 1). This concept is known as mitohormesis. It has been proposed that mitohormesis promotes improved

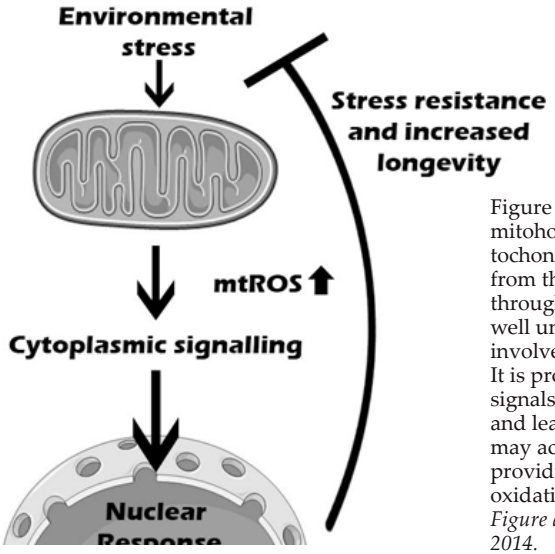


Figure 1: Simplified representation of mitohormesis. Perturbations of mitochondrial function are transmitted from the mitochondria to the cytosol through mechanisms which are not well understood but are thought to involve ROS as a signalling molecule. It is proposed that these cytoplasmic signals communicate with the nucleus and lead to a nuclear response which may activate cytoprotective pathways providing a level of resistance to mild oxidative stress.
Figure adapted from Yun and Finkel, 2014.

longevity and mitochondrial health through the increase of ROS production which may encourage long-lasting oxidative stress defences in the mitochondria.[18]

A recent study has found that muscle mitormesis prolongs lifespan through the repression of the insulin signalling pathway in *Drosophila*.[14] In this study *Drosophila* models exhibiting mitochondrial injury of muscular tissue were used to identify the signalling pathways involved in mitochondrial dysfunction. The study identified UPR^{mt} as a player in the age-dependant deterioration of muscle tissue along with the activation of insulin-like growth factor binding proteins (IGFBPs). These pathways associated with longevity were found promote mitophagy and mitochondrial quality control.

In mammals there are a number of IGFBPs which suggests that they may be secreted in human muscular tissues after mitochondrial injury.[2] The interaction between mitohormesis and these longevity pathways may have significant implications in the treatment of diseases associated with defective insulin signalling. A recent study examined *C.elegans* daf-2 mutants which are characterised by defects in insulin/IGF-1 signalling (iLLS).[19] The role of defective iLLS in increased longevity has already been described in previous studies.[14] It was found that the lifespan of these mutants increased by approximately 60% when compared to their wild-type counterparts.[19] These mutants show respiratory defects and increased ROS production. However, the ROS are thought to have activated ROS defence enzymes resulting in decreased levels of ROS.

A recent study has had surprising results, linking low levels of arsenite exposure to increased longevity in *C.elegans* through mitohormesis.[20] A very low dosage of arsenite was found to promote resistance against various mitochondrial perturbations while increasing the lifespan of the organism. As arsenite was observed to promote the production of ROS it was suggested that these ROS activate the arsenite in a toxin-mediated longevity pathway. Another study has demonstrated that mtROS may act as a signal for increased longevity in C. elegans through an intrinsic apoptosis pathway.[21] The study analysed *isp-1* and *nuo-6* ETC (electron transport chain) mutant organisms that were shown to exhibit increased mitochondrial dysfunction and mtROS production when compared to their wild-type counterparts. However, these mutants also demonstrated increased longevity. It was shown that the mtROS signal is only transmitted by the intrinsic apoptosis signalling pathway in the presence of CED-13, which accumulates in response to DNA damage (Figure 2.). The increased production of mtROS activates a distinctive gene expression pattern which promotes increased longevity in the organism. In the wild-type organism this mechanism may provide a level of protection to the nemotode when under stressful conditions.

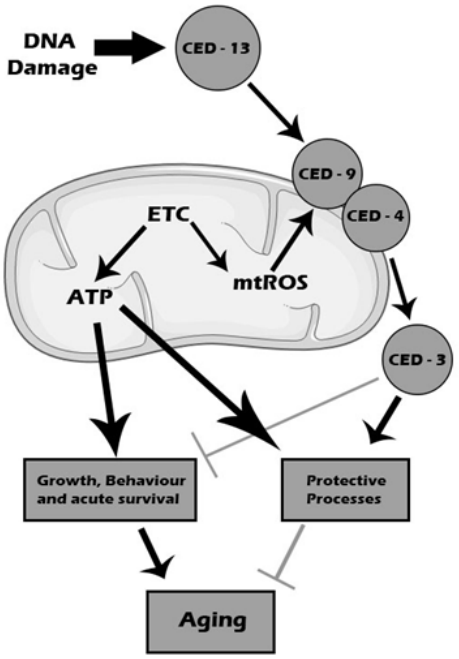


Figure 2: The intrinsic apoptosis pathway and mtROS signalling in *C.elegans*. This model demonstrates how the intrinsic apoptosis pathway regulates aging in *C.elegans* through mtROS signalling. The activation of the mtROS from the ETC is carried out by the alternative BH-3 protein CED-13. This induces the activation of the CED signalling pathway in the presence of increased levels of mtROS. This mechanism is proposed to provide protection against various perturbations to the mictochondria, inhibiting the aging process.
Figure adapted from Yee et al., 2014.

Conclusions

The mitochondria has played a central role in aging research over the past 50 years. The widely accepted hypothesis is that the accumulation of mitochondrial mutations is a principle force in driving the aging process. This hypothesis has been supported by a number of studies carried out thanks to the development of aging animal models. It was previously thought that the production of free radicals in the mitochondria damage mitochondrial DNA and contribute significantly to age-related mitochondrial dysfunction. However, this claim has been heavily refuted by a large number of recent studies providing data demonstrating that, in many cases, oxidative damage does not correlate with increased lifespan. Underlying pathways and key molecular players that directly impact lifespan have since been identified thanks to transgenic studies in mice, *Drosophila* and *C.elegans*. One such example is insulin/IGF-1 signalling which is inhibited during mitohormesis and promotes longevity in *Drosophila*. It has been proposed that ROS may induce cell senescence independently of oxidative stress and act as important signalling molecules in the poorly understood stress response mechanisms of mitohormesis and other stress-response pathway. It is these stress response mechanisms that have proven to be essential in our understanding of how the mitochondria regulates lifespan and aging. Evidence from multiple studies has suggested that moderate levels of mitochondrial perturbations activate these stress response mechanisms

and ultimately may provide long-term protection to the mitochondria. Therefore, identifying the underlying pathways of these responses may be fundamental in the development of therapeutic strategies to prevent mitochondrial-related disorders and age-related diseases.

References

1. Goldsmith, C.G. An Introduction to Biological Aging Theory. Crownsville, *MD: Azinet Press;* 2014. (2) 6-7.

2. Riera, CE., Merkwirth, C., De Magalhaes Filho C., Dillin, A. Signaling Networks Determining Lifespan. *Annu. Rev. Biochem*. 2016; 85: 35-64

3. Armanios, M., De Cabo R., Mannick., J., Partridge, L., Van Deursen, J., Villeda, S. Translational Strategies in Aging and Age-related Disease. *Nat Med*. 2015; 21(12): 1395-99.

4. Burgstaller, JP., Johnson, IG., and Poulton, J. Mitochondrial DNA Disease and Developmental Implications for Reproductive Strategies. *Mol. Hum. Reprod*. 2015; 21(1): 11-22.

5. Bratic, A., Larsson, NG. The Role of Mitochondria in Aging. *J Clin Invest*. 2013; 123(3): 951-7.

6. Ross, JM., Stewart, JB., Hagström, E., Brené, S., Mourier, A., Coppotelli, G., Freyer, C., Lagouge, M., Hoffer, BJ.,Olson, L., Larsson, NG. Germline Mitochondrial DNA Mutations Aggravate Ageing and Can Impair Brain Development. *Nature*. 2013; 501(7467): 412–15.

7. Edgar, DY., Wredenberg, A., Calvaruso, MA., Nijtmans, L., Nedergaard, J., Cannon, B.,Larsson, NG., Trifunovic, A. Random Point Mutations with Major Effects on Protein-coding Genes are the Driving Force behind Premature Aging in mtDNA Mutator Mice. *Cell Metab*. 2009; 10(2): 131–38.

8. Chial H, Craig J. mtDNA and Mitochondrial Diseases. *Nature Education.* 2008; 1(1): 217.

9. Larsson N-G (2010). Somatic Mitochondrial DNA Mutations in Mammalian Aging. *Annu. Rev. Biochem*. 2010; 79: 683–706.

10. Trifunovic, A., Wrendenberg, A., Falkenberg, M., Spelbrink, JN., Rovio, AT., Bruder, CE., Bohlooly, M., Gidlof, S., Oldfors, A., Wiborn, R., Tornell, J., Jacobs, HT., and Larsson N-G. Premature Ageing in Mice Expressing Defective Mitochondrial DNA Polymerase. *Nature.* 2004; 429(6990): 417-23.

11. Kujoth, GC., Hiona, A., Pugh, TD., Someya, S., Panzer, K., Wohlgemuth, SE., Hofer, T., Seo, AY., Sullivan, R., Joblin, WA., Morrow, JD., Remmen, HV., Sedivy, JM., Yamasoba, T., Tanokura, M., Weindruch, R., Leewenburgh, C., and Prolla, TA. Mitochondrial DNA Mutations, Oxidative Stress, and Apoptosis in Mammalian Aging. *Science*. 2005; 309(5733): 481-84.

12. Harman, D. Aging: A Theory based on Free Radical and Radiation Chemistry. *Journal of Gerontology.*1956; 11(3): 298–300

13. Tyynismaa, H., Mjosund, K., Wanroii, S., Lappalainen, I., Ylikallio, E., Jalanko, A., Spelbrink, J., Patau, A., Suomalainen, A. Mutant Mitochondrial Helicase Twinkle Causes Multiple mtDNA Deletions and a Late-onset Mitochondrial Disease in Mice. *PNAS*. 2005; 102(49): 17687-92.

14. Owusu-Ansah, E., Song, W., Perrimon, N. Muscle Mitohormesis Promotes Longevity *via* Systemic Repression of Insulin Signalling. *Cell*. 2013; 155(3): 699-712.

15. Hekimi, S., Lapointe, J., Wen, Y. Taking a "Good" Look at Free Radicals in the Aging Process. *Trends Cell Biol*. 2011; 21(10): 569-76.

16. Yang, W., Hekimi, S. A Mitochondrial Superoxide Signal triggers increased longevity in Caenorhabditis elegans. *PLoS Biol*.2010; 8(12).

17. Yun, J. Finkel, T. Mitohormesis. *Cell Metab.* 2014; 19(5): 757-66.

18. Ristow, M., Zarse, K. How Increased Oxidative Stress Promotes Longevity and Metabolic Health: The Concept of Mitochondrial Hormesis (Mitohormesis). *Exp Gerontol.* 2010; 45(6): 410-8.

19. Zarse, K., Schmeisser, S., Groth, M., Priebe, S., Beuster, G., Kulow, D., Guthke, R., Platzer, M., Kahn, CR., Ristow, M. Impaired Insulin/IGF1 Signaling Extends Lifespan by Promoting Mitochondrial L-Proline Catabolism to Induce a Transient ROS Signal. *Cell Metab.* 2012; 15(4): 451-65.

20. Schmeisser, S., Schmeisser, K., Weimer, S., Groth, M., Priebe, S., Fazius, E., Kuhlow, D., Prick, D., Einax, JW., Guthke, R., Platzer, M., Zarse, K., Ristow, M. Mitochondrial Hormesis Links Low-dose Arsenite Exposure to Lifespan Extension. *Aging Cell*. 2013; 12(3) 508-17.

21. Yee, C., Yang, W., Hekimi, S. The Intrinsic Apoptosis Pathway Mediates the Pro-longevity Response to Mitochondrial ROS in *C. elegans. Cell*. 2014; 157(4): 897-909.

22. K., Palikaras., E, Lionaki., N, Tavernarakis. Balancing Mitochondrial Biogenesis and Mitophagy to Maintain Energy Metabolism Homeostasis. *Nature.* 2015; 22(9): 1399-1401.

The Eye as a Window to Immunity:

Aetiological and Therapeutic Insights Gained from the Para-inflammatory Perspective in Age-related Macular Degeneration.

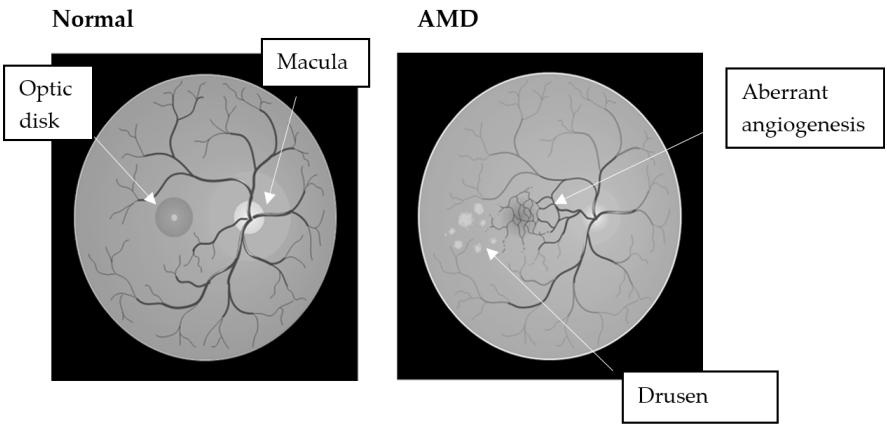
Ronan Treanor
Senior Sophister
Molecular Medicine

Age-related macular degeneration (AMD) is recognised by the World Health Organisation (WHO) as the leading cause of blindness in industrialised countries. Likewise, AMD is the most common cause of registered blindness in the Republic of Ireland, with an estimated prevalence of 7.2% in those aged 50 or older. An ageing population suggests a projected increase in the incidence of AMD, and as such, the disorder represents a pressing global issue. Clinically, late-stage AMD is divided into two forms: dry and wet (exudative). The clinical and diagnostic features of AMD are well noted, yet establishing a precise aetiology for both forms has proven challenging. At present, no therapy exists for the dry form of AMD. However, the success of anti-VEGF inhibitors in wet AMD demonstrates the utility of ongoing molecular research in this field. A number of modifiable (e.g. smoking) and unmodifiable risk factors (e.g. gene polymorphisms) are known to predispose to this disorder. As a consequence, the challenge of such research rests in deconstructing the multifactorial nature of AMD. Recent attention has been placed on the role of chronic, low-grade inflammation. This 'para-inflammatory' phenotype represents a shift in perspective from the classic view of inflammation arising from infection or serious injury and has been applied most notably in the context of metabolic disorders. Research into the inflammatory basis of AMD has yielded many potential key players in the progression of the disorder, including the complement system and the NLRP3 inflammasome. Attempts to understand the upstream events of this inflammatory phenotype suggest a fundamental failure within the aging retina, i.e. during the renewal of photoreceptor outer segments (POS). Moreover, this research demonstrates the rich complexity of the ocular immune system. Ultimately, targeting the key drivers of this inflammation will facilitate the ongoing design of molecular therapeutics for AMD.

Introduction

AMD is a progressive retinal disorder characterised by cellular death and dysfunction in the macula. The macula is located in the centre of the retina and contains a high-density of photoreceptors necessary for high-acuity vision – as a consequence, macular degeneration can result in central vision loss. AMD accounts for the majority of registered blindness in the Republic of Ireland and other industrialised nations.[1,2] AMD is identified in the clinical setting by the appearance of soft drusen deposits and lipofuscin (see Figure 1).[3] Drusen deposits are highly heterogeneous mixtures and consist of inflammation-related factors (e.g. complement proteins), lipid oxidation by-products (e.g. carboxyethylpyrrole (CEP)) and other cellular debris (amyloid-β).[4,5] In comparison, lipofuscin accumulates within the retinal pigment epithelium (RPE) as the result of a dysfunction in POS phagocytosis (i.e. incomplete digestion).[6] POS comprise the outer segments of photoreceptors (see Figure 1) and allow the photoreceptor (cone/rod) to absorb light. POS require daily renewal as a result of constant light exposure.[7] The RPE phagocytoses and recycles shed POS in a circadian fashion through a lysosomal-mediated process known as 'heterophagy'.[8]. *Hetero*phagy describes the ingestion and digestion of exogenous material by a cell (as opposed to autophagy). Recycling of POS replenishes retinoid (chromophore) levels in the photoreceptors and is known to aid the survival of the RPE.[9]

Following diagnosis, AMD is categorised into two clinical forms: dry and wet (exudative) – see Figure 1. The wet form is responsible for the majority of severe central vision loss attributable to AMD. Therapies for the wet form of AMD aim to either prevent or remove abnormal 'leaky' blood vessels developing in the retina.[10] At present, no FDA-approved therapies exist for the dry form of AMD, although two large clinical trials (AREDS, AREDS2) have demonstrated a possible benefit of antioxidant supplementation.[11,12]



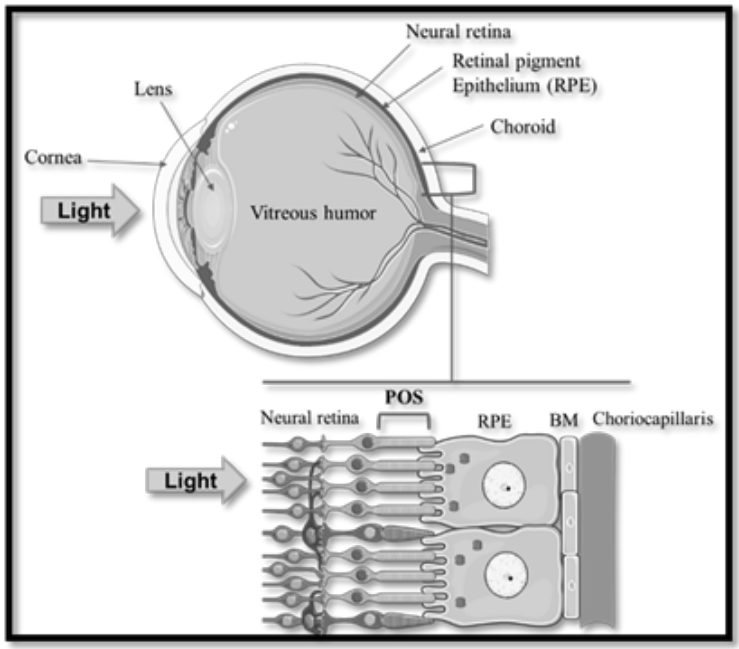


Figure 1: First panels feature a healthy retina and a retina showing clinical features of wet AMD. Drusen deposition is an early clinical hallmark alongside lipofuscin accumulation.[3] Drusen is typically found between the RPE and BM, whereas lipofuscin is found within the RPE.[3] Early AMD can then progress to either dry or wet AMD. Dry AMD is characterised by photoreceptor and RPE atrophy. Wet AMD is characterised by aberrant angiogenesis emanating from the choroid. The bottom panel outlines the basic anatomy of the retina. The RPE cells form a monolayer in the outer retina and are intimately associated with the photoreceptors.
Abbreviations used: RPE: retinal pigment epithelium; POS: photoreceptor outer segment; BM: Bruch's membrane. Adapted from Klettner[13]
[All figures in this essay are drawn using templates from Servier Medical Art Images]

The lack of therapies for dry AMD reflect current uncertainty about the progression of the disorder. Environmental oxidative stress (e.g. smoking) is recognised as a predisposing factor and is reflected in the presence of lipid oxidation by-products (e.g. malondialdehyde (MDA)) in drusen.[14] Incidentally, the RPE is a constitutive producer of complement (regulatory) factor H (CFH), which acts as a pattern-recognition receptor (PRR) for MDA and C-reactive protein.[15] The positive correlation between AMD incidence and the Y402H CFH gene variant – which affects binding to CRP - highlights the crucial role of controlling inflammation in the retina.[16]

Unfortunately, the aging, AMD-afflicted retina does not appear to deal sufficiently with this chronic oxidative stress. This failure is believed to be a precipitating factor for the chronic, low-grade sterile inflammation recognised in AMD.[17] The evidence

for this inflammatory phenotype comes from a variety of sources, as presented in Table 1. Various components of the innate immune system are noted, including complement and the NLRP3 inflammasome. Intriguingly, such inflammation arises in the absence of an infection, autoimmune reaction or obvious tissue injury and is strikingly reminiscent of the inflammation seen in obesity and metabolic disorders.[18] Recognising this ''para-inflammatory'' phenotype has provided both aetiological and therapeutic insights into AMD. This review will briefly outline the concept of 'para-inflammation' in the context of AMD, describe the links between defective heterophagy and downstream inflammatory consequences, as well as draw attention to recent therapeutic developments in the treatment of AMD.

Immune Component	Observation(s)	References
-	No infectious agent identified	-
Complement System	CFH SNP association; activated complement components found in drusen.	16, 19, 20
NLRP3 Inflammasome	Upregulation in the RPE of AMD patients; therapeutic potential versus NLRP3-derived cytokines IL-1β and IL-18.	21-23
Drusen Composition	Various inflammation-related factors.	4
Microglial Activation	CXCR1 SNP association with AMD in some populations. Cxcr1-/- mice exhibit RPE degeneration.	24-27

Table 1: Evidence for inflammation in AMD.
Abbreviations not mentioned previous: SNP: single nucleotide polymorphism; CXCR1: C-X-C motif chemokine receptor 1.

Para-inflammation and Heterophagy: a Defect in Routine Operations?

Ruslan Medzhitov proposed the term *para-inflammation* in 2008 to describe the intermediate inflammatory state nestled between the extremes of tissue infection/injury and basal homeostasis (see Figure 2).[19] This inflammatory phenotype, he suggests, arises from mild tissue malfunction/stress and ultimately represents an attempt to restore ''tissue functionality and homeostasis''. However, if such conditions persist, this low-grade inflammation can become chronic and eventually ''maladaptive''.

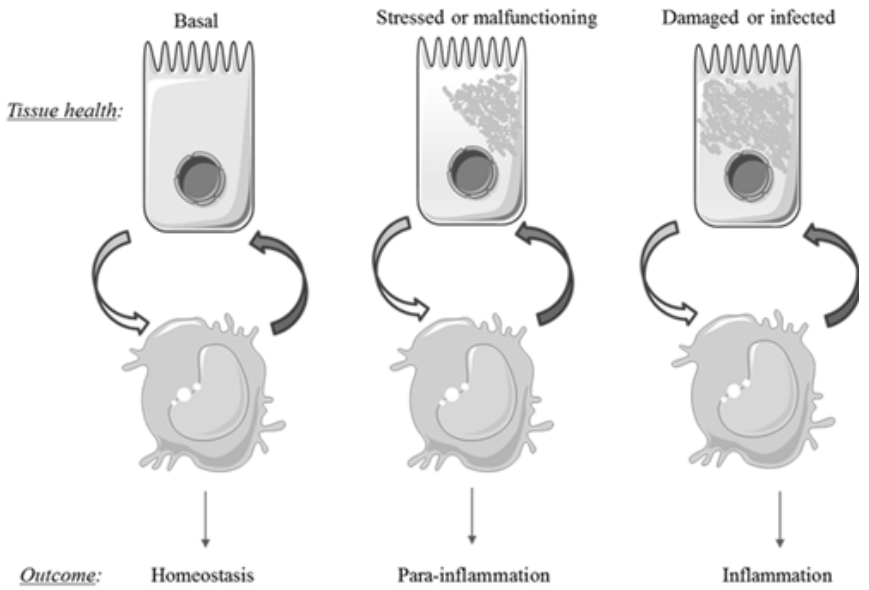


Figure 2: Para-inflammation is proposed to rest along a continuum nestled between the extremes of tissue homeostasis and outright inflammation. The degree of shading indicates the extent of ill-health. This model also proposes that tissue-resident macrophages (TRM) play a crucial role in the surveillance of tissue health.
Adapted from Medzhitov[19]

Chronic para-inflammation provides a useful conceptual framework for understanding the inflammatory retinal phenotype in AMD. This is most evident when considering the intimate anatomical relationship between the photoreceptors and the RPE (see Figure 1). The photoreceptors rely on the RPE to replenish retinoid levels by lysosomal-mediated phagocytosis of shed POS, whereas the RPE benefits from the concomitant synthesis of the anti-apoptotic molecule, neuroprotection D1.[9] However, this relationship is not strictly maintained over time as evidenced by the positive association between lipofuscin accumulation and ageing.[3] In the context of chronic oxidative stress (e.g. photo-oxidation) incomplete lysosomal degradation of POS can lead to lipid oxidation end-products such as CEP and MDA.[14] Both products are inherently antigenic - a CEP-albumin adduct alone is sufficient to elicit an immunoglobulin response and a dry-AMD pathology in mice, whereas MDA is sequestered by CFH (constitutively produced by the RPE).[20,21] The build-up of lipofuscin may also contribute to drusen formation beneath the RPE.[22] Isolated drusen is known to directly activate the NLRP3 inflammasome and so could further contribute to the already burgeoning inflammatory milieu.[23]

Ultimately, initial attempts by the RPE to adapt to mild tissue stress could generate an acute para-inflammatory response, which, under persistent oxidative

stress (oxidised POS) can become chronic. If combined with the appropriate environmental (smoking) and genetic (CFH gene SNP) conditions, this *chronic para-inflammatory* state may become dysregulated (maladaptive) and eventually accelerate the onset of macular degeneration.

Thus, applying Medzhitov's concept is useful in understanding the general inflammatory phenotype of AMD, particularly in light of defective lysosomal-mediated POS phagocytosis. Moreover, links between a defective lysosomal system and activation of the NLRP3 inflammasome have been established in the literature.[24] Such links suggest the NLRP3 inflammasome as an important sensor of the perturbed conditions in the RPE.

A Defective Lysosomal System and NLRP3 Activation in the RPE

The NLRP3 inflammasome is an innate immune sensor which acts as a molecular platform for caspase-1-mediated cleavage of pro-IL-1β/IL-18 into mature cytokines. The release of cytokines allows for communication with the immune system at-large and triggers either a pro-inflammatory or anti-inflammatory response.[25] The NLRP3 inflammasome consists of three major constituents: a sensor (NLRP3 protein), an adaptor protein (ASC) and an effector molecule (caspase-1) - the term 'inflammasome' describes the multimeric protein complex arising from the assembly of these major components. Induction of the NLRP3 inflammasome requires two steps: 1. Priming signal e.g. LPS-TLR4 and 2. An activation signal e.g. mitochondrial ROS.[26,27] The NLRP3 inflammasome is activated in response to various PAMPs and DAMPs and uncontrolled activation is associated with aberrant chronic inflammation.[18,28]

The NLRP3 inflammasome was first implicated in AMD in 2012 by various sources.[23,29-31] It was shown to be upregulated in the RPE and drusen of human AMD donors versus unaffected controls[29,32] In vitro activation of the NLRP3 inflammasome has also been demonstrated with common lipid peroxidation products found in drusen (i.e. CEP and 4-hydroxynonenal).[23,30] Destabilisation of the lysosomal system has been strongly implicated in these studies – direct lysosomal permeabilisation by the lysosomotropic agent Leu-Leu-OMe induced caspase-1 activation and IL-1β secretion in IL-1α primed ARPE-19 cells.[29] Moreover, activation of the NLRP3 inflammasome in this manner induced a form of caspase-1-dependent cell death known as 'pyroptosis'. The release of the lysosomal cysteine protease cathepsin B plays a key role in this process.[23,29] A2E, a constituent of lysosomal lipofuscin, directly disintegrates the membranes of intact lysosomes isolated from RPE cell lines and induces NLRP3 inflammasome-dependent IL-1β production upon stimulating ARPE-19 cells.[33,34] Furthermore, (blue-light) irradiation of lipofuscin in ARPE-19 and primary human RPE lead to lysosomal leakage, caspase-1 activation and IL-1β plus IL-18 secretion. Inhibition of cathepsin

B reduced the extent of cytokine secretion.[35] Although predominantly performed on cell lines in vitro, these findings exhibit the link between lysosomal disruption and NLRP3 inflammasome activation in RPE cells. As well as this, these studies show how lipofuscin (i.e. incompletely degraded POS) can play a role in lysosomal destabilisation and NLRP3 inflammasome activation.

It should be noted that POS heterophagy is not the only lysosomal-mediated process occurring in the RPE. Although the RPE cell ingests and digests external material (e.g. POS heterophagy), it also undergoes the fundamental process of autophagy (internal digestion). Autophagy – which is necessary for the degradation of damaged cellular organelles and for the maintenance of energy homeostasis – has also been implicated in AMD progression. Analyses of human RPE samples revealed that autophagosome formation and levels of autophagy-associated proteins (ATG7, ATG9) increased in *healthy* old RPE versus *healthy* young RPE – however, autophagosome formation was significantly reduced in late-stage AMD-affected RPE cells versus healthy, age-matched controls.[36] These findings suggest that although autophagy may increase with age to tolerate enhanced oxidative stress/organelle dysfunction in *healthy* human RPE, a decline may result in the RPE of those with late-stage AMD. Moreover, autophagy has been shown to regulate the degradation of ubiquitinated NLRP3 inflammasomes in macrophages, which hints at a connection with NLRP3 activation in AMD.[37,38]

Therapeutic Prospects for AMD

The NLRP3 inflammasome represents a rich avenue for therapeutic exploitation – in particular *via* the associated cytokines, IL-1β and IL-18.[39] The role of IL-18 in wet AMD is controversial due to conflicting results and may contribute to *Alu* RNA induced RPE degeneration.[23,31,40] Nevertheless, preclinical trials in murine and nonhuman primate models suggest a potential role for recombinant IL-18 as an adjunctive immunotherapy alongside current VEGF-A inhibitors such as ranibizumab (Lucentis ®).[41,42] The proangiogenic role of IL-1β however, is certain and targeting IL-1β signalling has yielded promising results in murine models of wet AMD.[43,44] Novel therapeutic creations for wet AMD remain in demand given the need for more effective, less invasive and less frequent treatments – a selection of potential therapeutics are listed in Table 2.[45]

AMD	Name	Target	Phase	Company
Wet	Human recombinant IL-18	VEGFR2	Completed pre-clinical trial[41]	GlaxoSmithKline/ Trinity College Dublin
	Abicipar Pegol	Anti-VEGF and anti-PDGF	III[46]	Allergan
	Pan-90806 Tyrosine kinase inhibitor	VEGFR2	II[47]	PanOptica
	RetinoStat®	Lentiviral vector delivering endostatin and angiostatin	Completed Phase I[48]	Oxford BioMedica
Dry	Lampalizumab	Anti-complement factor D	III (see text).	Genetech/Roche
	Fenretinide	Serum retinal-binding protein	Completed Phase II [49]	ReVision Therapeutics
	Zumira®	Anti-C5	II/III[50]	Ophthotech
	LFG316	Anti-C5	Completed Phase II (however ineffective)[50]	Novartis

Table 2: Examples of recently developed therapeutics for the treatment of AMD (ongoing or recently completed clinical trials).

Novel therapeutics are also in the pipeline for dry AMD. A prominent target is the complement system, the aberrant activation of which is strongly implicated in AMD progression (see Table 2).[51] Complement proteins are a series of plasma-soluble 'first line of defence' mediators of the innate immune system. Numerous complement proteins have been found in drusen and gene variants show a significant association with AMD – thus complement may contribute to the para-inflammation in the AMD-afflicted retina.[15,16] One of the promising breakthroughs is lampalizumab, a single-affinity matured antigen-binding fragment (Fab) which inhibits the activation of C3 convertase, an important rate-limiting step in complement activation.[53] MAHALO, a recently completed phase II clinical trial involving lapalizumab, demonstrated significant reduction in geographic atrophy versus sham controls, particularly in those with a CFI SNP.[52,54,55] Currently, three

double-blinded, multicentre phase III clinical trials (OMASPECT, SPECTRI and CHROMA) are recruiting to demonstrate the efficacy, tolerability and safety of lampalizumab in a large cohort of advanced AMD patients.[56-58]

Conclusions

Age-related macular degeneration bears a significant burden in both the personal and societal context and is expected to present a major problem in the future, particularly given current demographics. Knowledge of the underlying inflammatory burden in AMD has unveiled an assortment of viable targets, including the NLRP3 inflammasome, IL-18 and the complement system. This information is critical for the establishment of a precise aetiology and the creation of appropriate therapeutics. The current lack of FDA-approved therapies for dry AMD highlights the need for such knowledge. Conversely, the impact of such knowledge is evident given ongoing clinical trials.

However, the upstream factors contributing to the para-inflammatory phenotype in AMD are not completely understood - albeit there exists evidence suggesting a role for a defective lysosomal system in the ageing RPE. As a consequence, this information gap represents a further series of questions for current research to investigate.

Acknowledgements

The author would like to thank Dr. S. Doyle and Dr. R. Mahony for their support throughout the year. The author would also like to thank Ms. Hughes and Mr. Schofield for their suggested edits. This work was adapted from the Molecular Medicine Research thesis ''To characterise the inflammatory and metabolic phenotype of the retinal pigment epithelium in response to oxidised photoreceptor outer segments'' for module BI4395. Figures were created using templates from the Servier Medical Art image bank.

References

1. World Health Organisation. Priority eye diseases.. Available from:. http://www.who.int/blindness/causes/priority/en/.

2. Green D, Ducorroy G, McElnea E, Naughton A, Skelly A, O'Neill C, *et al*. The Cost of Blindness in the Republic of Ireland 2010-2020. *J Ophthalmol*. 2016;2016:4691276.

3. de Jong P. Mechanisms of disease: Age-related macular degeneration. *New England Journal of Medicine*. 2006;355(14):1474-85.

4. Anderson DH, Mullins RF, Hageman GS, Johnson LV. A role for local inflammation in the formation of drusen in the aging eye. *Am J Ophthalmol*. 2002;134(3):411-31.

5. Anderson DH, Talaga KC, Rivest AJ, Barron E, Hageman GS, Johnson LV. Characterization of beta amyloid assemblies in drusen: the deposits associated with aging and age-related macular degeneration. *Exp Eye Res.* 2004;78(2):243-56.

6. Kennedy CJ, Rakoczy PE, Constable IJ. Lipofuscin of the retinal pigment epithelium: a review. *Eye (Lond).* 1995;9 (Pt 6):763-71.

7. Baker SA, Kerov V. Photoreceptor inner and outer segments. *Curr Top Membr.* 2013;72:231-65.

8. Lei L, Tzekov R, McDowell JH, Smith WC, Tang S, Kaushal S. Formation of lipofuscin-like material in the RPE Cell by different components of rod outer segments. *Exp Eye Res*. 2013;112:57-67.

9. Mukherjee PK, Marcheselli VL, de Rivero Vaccari JC, Gordon WC, Jackson FE, Bazan NG. Photoreceptor outer segment phagocytosis attenuates oxidative stress-induced apoptosis with concomitant neuroprotectin D1 synthesis. *Proc Natl Acad Sci U S A.* 2007;104(32):13158-63.

10. Nowak JZ. AMD--the retinal disease with an unprecised etiopathogenesis: in search of effective therapeutics. *Acta Pol Pharm.* 2014;71(6):900-16.

11. Lutein + zeaxanthin and omega-3 fatty acids for age-related macular degeneration: the Age-Related Eye Disease Study 2 (AREDS2) randomized clinical trial. *Jama*. 2013;309(19):2005-15.

12.A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol.* 2001;119(10):1417-36.

13. Klettner A, Kauppinen A, Blasiak J, Roider J, Salminen A, Kaarniranta K. Cellular and molecular mechanisms of age-related macular degeneration: from impaired autophagy to neovascularization. *Int J Biochem Cell Biol.* 2013;45(7):1457-67.

14. Handa JT. How does the macula protect itself from oxidative stress? *Mol Aspects Med.* 2012;33(4):418-35.

15. Chen M, Forrester JV, Xu H. Synthesis of complement factor H by retinal pigment epithelial cells is down-regulated by oxidized photoreceptor outer segments. *Exp Eye Res.* 2007;84(4):635-45.

16. Fritsche LG, Igl W, Bailey JN, Grassmann F, Sengupta S, Bragg-Gresham JL, *et al*. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat Genet.* 2016;48(2):134-43.

17. Kauppinen A, Paterno JJ, Blasiak J, Salminen A, Kaarniranta K. Inflammation and its role in age-related macular degeneration. *Cell Mol Life Sci.* 2016;73(9):1765-86.

18. Wen H, Ting JP, O'Neill LA. A role for the NLRP3 inflammasome in metabolic diseases--did Warburg miss inflammation? *Nat Immunol.* 2012;13(4):352-7.

19. Medzhitov R. Origin and physiological roles of inflammation. *Nature.* 2008;454(7203):428-35.

20. Hollyfield JG, Bonilha VL, Rayborn ME, Yang X, Shadrach KG, Lu L, *et al*. Oxidative damage-induced inflammation initiates age-related macular degeneration. *Nat Med.* 2008;14(2):194-8.

21. Weismann D, Hartvigsen K, Lauer N, Bennett KL, Scholl HP, Charbel Issa P, *et al*. Complement factor H binds malondialdehyde epitopes and protects from oxidative stress. *Nature.* 2011;478(7367):76-81.

22. Zhou J, Jang YP, Kim SR, Sparrow JR. Complement activation by photooxidation products of A2E, a lipofuscin constituent of the retinal pigment epithelium. *Proc Natl Acad Sci U S A*. 2006;103(44):16182-7.

23. Doyle SL, Campbell M, Ozaki E, Salomon RG, Mori A, Kenna PF, *et al*. NLRP3 has a protective role in age-related macular degeneration through the induction of IL-18 by drusen components. *Nat Med.* 2012;18(5):791-8.

24. Kaarniranta K, Sinha D, Blasiak J, Kauppinen A, Vereb Z, Salminen A, *et al*. Autophagy and heterophagy dysregulation leads to retinal pigment epithelium dysfunction and development of age-related macular degeneration. *Autophagy.* 2013;9(7):973-84.

25. Turner MD, Nedjai B, Hurst T, Pennington DJ. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochim Biophys Acta.* 2014;1843(11):2563-82.

26. Misawa T, Takahama M, Kozaki T, Lee H, Zou J, Saitoh T, *et al*. Microtubule-driven spatial arrangement of mitochondria promotes activation of the NLRP3 inflammasome. *Nat Immunol.* 2013;14(5):454-60.

27. Bauernfeind FG, Horvath G, Stutz A, Alnemri ES, MacDonald K, Speert D, *et al*. Cutting edge: NF-kappaB activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *J Immunol.* 2009;183(2):787-91.

28. Broderick L, De Nardo D, Franklin BS, Hoffman HM, Latz E. The inflammasomes and autoinflammatory syndromes. *Annu Rev Pathol.* 2015;10:395-424.

29. Tseng WA, Thein T, Kinnunen K, Lashkari K, Gregory MS, D'Amore PA, *et al*. NLRP3 inflammasome activation in retinal pigment epithelial cells by lysosomal destabilization: implications for age-related macular degeneration. *Invest Ophthalmol Vis Sci.* 2013;54(1):110-20.

30. Kauppinen A, Niskanen H, Suuronen T, Kinnunen K, Salminen A, Kaarniranta K. Oxidative stress activates NLRP3 inflammasomes in ARPE-19 cells--implications for age-related macular degeneration (AMD). *Immunol Lett.* 2012;147(1-2):29-33.

31. Tarallo V, Hirano Y, Gelfand BD, Dridi S, Kerur N, Kim Y, *et al*. DICER1 loss and Alu RNA induce age-related macular degeneration *via* the NLRP3 inflammasome and MyD88. *Cell.* 2012;149(4):847-59.

32. Wang Y, Hanus JW, Abu-Asab MS, Shen D, Ogilvy A, Ou J, *et al*. NLRP3 Upregulation in Retinal Pigment Epithelium in Age-Related Macular Degeneration. *Int J Mol Sci.* 2016;17(1).

33. Schutt F, Bergmann M, Holz FG, Kopitz J. Isolation of intact lysosomes from human RPE cells and effects of A2-E on the integrity of the lysosomal and other cellular membranes. *Graefes Arch Clin Exp Ophthalmol*. 2002;240(12):983-8.

34. Anderson OA, Finkelstein A, Shima DT. A2E induces IL-1ss production in retinal pigment epithelial cells *via* the NLRP3 inflammasome. *PLoS One*. 2013;8(6):e67263.

35. Brandstetter C, Mohr LK, Latz E, Holz FG, Krohne TU. Light induces NLRP3 inflammasome activation in retinal pigment epithelial cells *via* lipofuscin-mediated photooxidative damage. *J Mol Med* (Berl). 2015;93(8):905-16.

36. Mitter SK, Song C, Qi X, Mao H, Rao H, Akin D, *et al*. Dysregulated autophagy in the RPE is associated with increased susceptibility to oxidative stress and AMD. *Autophagy.* 2014;10(11):1989-2005.

37. Shi CS, Shenderov K, Huang NN, Kabat J, Abu-Asab M, Fitzgerald KA, *et al*. Activation of autophagy by inflammatory signals limits IL-1beta production by targeting ubiquitinated inflammasomes for destruction. *Nat Immunol*. 2012;13(3):255-63.

38. Liu RT, Wang A, To E, Gao J, Cao S, Cui JZ, *et al*. Vinpocetine inhibits amyloid-beta induced activation of NF-kappaB, NLRP3 inflammasome and cytokine production in retinal pigment epithelial cells. *Exp Eye Res.* 2014;127:49-58.

39. Ozaki E, Campbell M, Doyle SL. Targeting the NLRP3 inflammasome in chronic inflammatory diseases: current perspectives. *J Inflamm Res.* 2015;8:15-27.

40. Hirano Y, Yasuma T, Mizutani T, Fowler BJ, Tarallo V, Yasuma R, *et al*. IL-18 is not therapeutic for neovascular age-related macular degeneration. *Nat Med.* 20. United States2014. p. 1372-5.

41. Doyle SL, Lopez FJ, Celkova L, Brennan K, Mulfaul K, Ozaki E, *et al*. IL-18 Immunotherapy for Neovascular AMD: Tolerability and Efficacy in Nonhuman Primates. *Invest Ophthalmol Vis Sci.* 2015;56(9):5424-30.

42. Doyle SL, Ozaki E, Brennan K, Humphries MM, Mulfaul K, Keaney J, *et al*. IL-18 attenuates experimental choroidal neovascularization as a potential therapy for wet age-related macular degeneration. *Sci Transl Med*. 2014;6(230):230ra44.

43. Olson JL, Courtney RJ, Rouhani B, Mandava N, Dinarello CA. Intravitreal anakinra inhibits choroidal neovascular membrane growth in a rat model. *Ocul Immunol Inflamm.* 2009;17(3):195-200.

44. Marneros AG. NLRP3 inflammasome blockade inhibits VEGF-A-induced age-related macular degeneration. *Cell Rep.* 2013;4(5):945-58.

45. Rofagha S, Bhisitkul RB, Boyer DS, Sadda SR, Zhang K. Seven-year outcomes in ranibizumab-treated patients in ANCHOR, MARINA, and HORIZON: a multicenter cohort study (SEVEN-UP). *Ophthalmology.* 2013;120(11):2292-9.

46. ClinicalTrials.gov. Safety and Efficacy of Abicipar Pegol in Patients With Neovascular Age-related Macular Degeneration. Available from: https://clinicaltrials.gov/ct2/show/NCT02462486;

47. PanOptica. PanOptica Reports Positive Results from Phase 1/2 Clinical Trial of PAN-90806, a Novel Topical Anti-VEGF Eye Drop Available from: http://www.chemdiv.com/panoptica-reports-positive-results-phase-12-clinical-trial-pan-90806-novel-topical-anti-vegf-eye-drop/;

48. Campochiaro PA, Lauer AK, Sohn EH, Mir TA, Naylor S, Anderton MC, *et al*. Lentiviral Vector Gene Transfer of Endostatin/Angiostatin for Macular Degeneration (GEM) *Study. Hum Gene Ther.* 2017;28(1):99-111.

49. Mata NL, Lichter JB, Vogel R, Han Y, Bui TV, Singerman LJ. Investigation of oral fenretinide for treatment of geographic atrophy in age-related macular degeneration. *Retina.* 2013;33(3):498-507.

50. ClinicalTrials.gov. Intravitreal LFG316 in Patients With Age-related Macular Degeneration (AMD) Available from: https://clinicaltrials.gov/ct2/show/NCT01527500?term=LFG-316&rank=1;

51. Anderson DH, Radeke MJ, Gallo NB, Chapin EA, Johnson PT, Curletti CR, *et al*. The pivotal role of the complement system in aging and age-related macular degeneration: hypothesis re-visited. *Prog Retin Eye Res.* 2010;29(2):95-112.

52. Dunkelberger JR, Song WC. Complement and its role in innate and adaptive immune responses. *Cell Res.* 2010;20(1):34-50.

53. Le KN, Gibiansky L, Good J, Davancaze T, van Lookeren Campagne M, Loyet KM, *et al*. A mechanistic pharmacokinetic/pharmacodynamic model of factor D inhibition in cynomolgus monkeys by lampalizumab for the treatment of geographic atrophy. *J Pharmacol Exp Ther.* 2015;355(2):288-96.

54. Ehmann D. Complement Inhibition For the Treatment of GA. Available from:. https://www.reviewofophthalmology.com/article/complement-inhibition-for-the-treatment-of-ga.

55. Roche. Investor Update. Available from:. http://www.roche.com/investors/updates/inv-update-2013-08-27.htm.

56. ClinicalTrials.gov. CHROMA. Available from: https://clinicaltrials.gov/ct2/show/NCT02247479?term=CHROMA&rank=1

57. ClinicalTrials.gov. SPECTRI. Available from: https://clinicaltrials.gov/ct2/show/NCT02247531?term=spectri&rank=1

58. ClinicalTrials.gov. OMASPECT. Available from: https://clinicaltrials.gov/ct2/show/NCT02745119?term=OMASPECT&rank=1 [heterochromatin formation. Genes and Development, 16, 479–489

Horizontal Gene Transfer in Modern Plants:

Detection, Evaluation and Modern Manipulation

Conor Rossi
Junior Sophister
Genetics

Horizontal Gene Transfer (HGT), also referred to as Lateral Gene Transfer (LGT), is the process by which genetic information moves across barriers of distinct organisms and becomes integrated into a new host genome. It distinguishes itself from vertical transmission, which is the transfer of genetic material from parent to offspring through reproduction. The reaches of HGT today extend from natural novel gene acquisition to artificial manipulation of genome. A growing body of evidence is proving that HGT is ubiquitous in nature. In this review, I discuss the importance of this mechanism of gene transfer in the Viridiplantae clade, as described by Leliaert et al.[1] This simplified group of plastid-containing organisms includes flowering plants, ferns, mosses and green algae. I evaluate the barriers of HGT, different mechanisms of HGT in these species and discuss artificial manipulation of HGT in modern plant species.

Introduction

HGT was first observed in 1928 in strains of the eubacterial species *Streptococcus pneumoniae* by Frederick Griffith.[2] Griffith observed an acquisition of virulence by a non-virulent strain of *S. pneumoniae* when cultured with heat-killed virulent strains. The non-virulent strain was able to uptake free DNA from dead cells in solution and acquire novel virulent gene function. This was later discovered to be caused by plasmid DNA in a process termed '*transformation*' in a landmark experiment by Avery, MacLeod and McCarty in 1944.[3] Although originally disputed by the scientific community, transformation is now a well-established method of gene acquisition and evolution for bacteria with approximately 80 species displaying natural competence to accept exogeneous DNA.[4] In 1959, the first instance of gene flow to different species was recorded in bacteria by Akiba and Ochia.[5] This showed that an apparent species boundary does not prevent gene transfer. The acquisition

of antibiotic resistance from a different species suggest that HGT may admonish acquisition of new genes. These genes may assist in adaptation to rapidly changing environments. Further work in the 1970s by Fritz Went and Krasilov suggested that eukaryotic gene transfer may confer similar traits through their work on flowering plants.[6] This marked the start of an appreciation of the importance of HGT in gene acquisition for eukaryotes. The acquisition of novel genes by HGT has been since shown to be a driver in eukaryotic evolution.[6] Comparative studies of species and the genome sequencing revolution have implicated HGT in the complicated phylogenies of many plant species. However, significant barriers exist for the detection of such HGT events.

Inference of Horizontal Gene Transfer

The first recorded instance of HGT was based on phenotypic evidence of antibiotic resistance but most contemporary methods of inferring gene transfer are not phenotypic. Instead, contemporary methods focus on phylogenetic and parametric genomic differences.[7] Phylogenetic methods evaluate the evolutionary history of genes by identifying conflicting genomic signatures like nucleotide composition, oligonucleotide frequencies or structural features of the genome.[8] These methods take advantage of the recent influx of sequenced genomes. By computationally comparing genomes, inferences of HGT may be made with a statistical context. However, they are not wholly reliable for detecting HGT as their ability to detect differences between sister taxa is limited by the extent of genomic divergence. Parametric difference focus on '*surrogate*' or non-tree values of genes like GC content.[9] Parametric analysis frequently operates under the assumption that the GC content is constant in genomes of species and so foreign gene inserts are detectable by the presence of DNA sequences with significantly different GC content to the rest of the host genome.[10] However, these assumptions are not particularly robust. For example, GC content is lower close to the replication terminus and higher in regions of highly expressed genes.[11,12] This may lead to endogenous sequences mistakenly interpreted as exogeneous. Parametric analysis also suffers from considerable bias from the assumption that exogeneous genes have different characteristics to endogenous genes.[13] Once a foreign piece of DNA has been inserted, it is also 'ameliorated' to reflect the general base composition of the host genome which may also obscure HGT genes.[10] Clearly, at present neither inference methods are wholly reliable and combination of methods should be used when evaluating cases of HGT.

Variability of Horizontal Gene Transfer in Plant Species

The importance of HGT in eukaryotic evolution has been questioned due to several factors. Firstly, the occurrence of HGT in eukaryotes varies in different lineages and mostly occurs in genomes like those of the understudied

unicellular eukaryotes, protists.[15] Often, well studied multicellular eukaryotes with segregated germ lines are more immune to HGT due to a less meaningful exposure to foreign DNA. For foreign genes to become heritable in a multicellular organism, these genes must be inserted into the germline (e.g. meristem of a plant) as these are the cells that will transfer genes that become system-wide in ancestors. A gene integrated into host DNA in a somatic cell will not be able to be passed to future generations.[14] Unicellular organisms circumvent this issue by their means of asexual reproduction (e.g. budding in yeasts) allowing all genes inherited by the daughter cells. A segregated germline is not the sole barrier to meaningful exposure to potential gene transfers. In a concept aptly named 'You Are What You Eat', Doolittle suggests that the proportion of genes horizontally transferred to the nucleus from bacteria to eukaryotes should be increased if a eukaryote predates them.[16] Indeed, it has been reported that 1-4% of the unicellular phagocytic protists' gene collection has bacterial HGT origins.[16] Phagocytosis is highly important for the integration of foreign genes as DNA is exposed close to the nucleus of the phagocytic cell. As a result non-phagocytic plant species have less instances of HGT even when unicellular. For example, the non-phagotrophic green alga *Chlamydomonas reinhardtii* has been proven to not harbour any bacterial genes.[14] Although there are significant barriers, there is sufficient prevalence of HGT in plants to consider its importance. Indeed, isolated events of HGT in plants have been proven to have significant influence on genome evolution and novel gene acquisition.[17]

Novel Gene Acquisition

Despite difficulty in detection, gene transfer has been determined to be the causative mechanisms of many selective advantages for plant evolution. For example, two genome studies on the plant species, *Physcomitrella patens* and *Arabidopsis thaliana*, uncovered 57 gene families acquired by bacterial genes.[18] These genes included xylem formation, defense, regulation of growth and the major biosynthetic pathway for auxin production. This suggests that HGT may have played an important role in transition of plants from an aquatic environment to land, a major evolutionary event.[18] This is an example of gene exaptation. Gene exaptation refers to a phenomenon of the development of fitness conferring genes after shifting from its original function.[19]

Plant mitochondria are proven to be transformation-competent whereas plastids do not share this functionality. A well studied case of fungi to angiosperm HGT is the mitochondrial Group I intron in the cox1 gene of *Peperomia polybotrya*.[20] This intron was found to be more closely related to fungal mitochondrial introns than introns of higher plants or green algae. Even more striking was the discovery that this HGT is reported to be caused by around 70 separate horizontal transfer events in 640 angiosperm genera.[20] It is suggested that after an original fungi to angiosperm HGT event, angiosperm-to-angiosperm HGT persisted dozens of

times.[21] The intron contains an open reading frame (ORF) that is posited to encode a site-specific endonuclease involved in intron mobility. This gene exaptation must have conferred some kind of positive fitness by evidence of its fixation and reinfection in angiosperms. Plant mitochondrial genomes are also repopulated with genes previously lost *via* transfer to nucleus in what's known as recapture HGT.[16] In an unprecedented example on recapture DNA, enormous transfers of 3.9 Mb of DNA was found in Amborella.[22] It contains six genomes equivalents of foreign mitochondrial DNA acquired from a range of species including green algae, mosses and fellow angiosperms - with some being whole genome transfers posited. A fusion-compatibility model has been proposed to explain these findings whereby *Amborella* captured whole mitochondria and fused them through genome recombination. This transfer has conferred many apparent benefits to the Amborella species including the acquisition of 61 genes from moss mtDNA. A further 753 kb of the *Amborella* mitochondrial genome was identified to be from other angiosperms. Of the 197 foreign genes captured, only 25% have intact ORF which suggests most are pseudogenes. However, the intact foreign DNA is thought to be a replacement of a missing *Amborella* mtDNA and has conferred some kind of selective advantage.[22]

Agrobacterium-Mediated Horizontal Gene Transfer

Agrobacterium rhizogenes and *Agrobacterium tumefaciens* are pathogenic bacteria that transfer plasmids into the genomes of host numerous plant genomes. They are a startling example of HGT in action. There is a conjugative transfer of DNA fragments (T-DNA) from the bacterium's tumour-inducing (Ti) plasmid or root-inducing (Ri) plasmid into the host cell.[23] Virulence genes in the Bacterium allows for the transit of T-DNA and their associated proteins to the host plant cell's nucleus where they may integrate and initiate gene expression. The single stranded T-DNA is converted to a double strand prior to integration. The host DNA breaks and DNA pathways repair and recombine the foreign DNA. During these processes, *Agrobacterium* has evolved to control and even utilize several pathways of host-plant defence response.[24] They express functional genes like that for the synthesis of opines which provides the invading *Agrobacterium* a selective advantage over most other bacteria since they cannot use opines as energy sources.[25] Other genes express for plant growth hormones, resulting in a large tumour where the cells can propagate in the 'crown gall tumors' or 'hairy root formations'. *Rhizobium etli* (a plant-symbiotic bacterium) possesses similar virulence genes.[26] Remarkably, sequences of *Agrobacterium* T-DNA have become reportedly fixed in the gene pools of *Nicotiana glauca*, a wild tobacco species and 291 accessions of sweet potato.[27,28] This suggests the parasitic T-DNA may also confer some selective advantage or be an example of neutral mutation and random fixation.

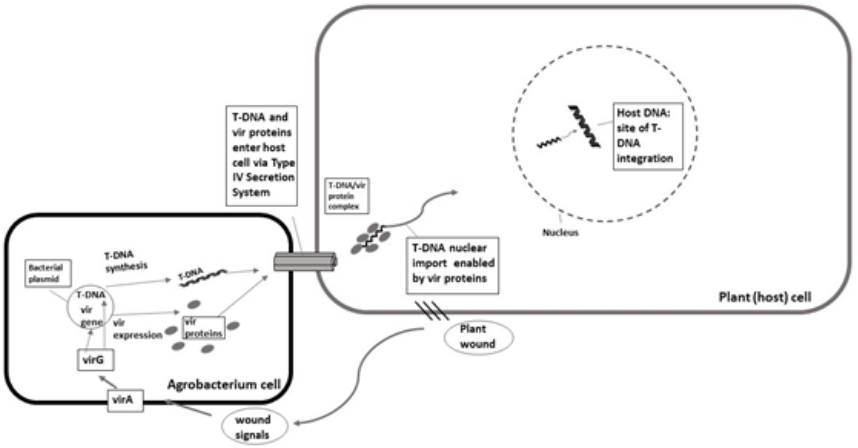


Figure 1: The mechanism of action of agrobacterium transformation. Plant signals from a plant wound triggers activation of virulence factors (VirA/G). VirA/G initiates T-DNA synthesis and vir gene expression. The T-DNA sequence and virulence factors import into the host plant cell *via* the bacterial type IV secretion system (T4SS). The virulent factors and T-DNA assemble in host cells. This assembly mediates the nuclear import of agrobacterium T-DNA which is then integrated into the host genome.
Adapted from Pitzschke & Hirt.[25]

Practical Application of HGT

The knowledge of HGT has led to many advances in biotechnology, none more so perhaps than genetic modification. Genetic modification can be broadly defined as a set of technologies used to alter the genetic makeup of an organism. One of these technologies is the artificial gene transfer horizontally across barriers of gene transfer. Extraneous DNA may be inserted into plants *via* Agrobacterium-mediated recombination.[29] As discussed above, *Agrobacterium* contains a tumour inducing (Ti) plasmid containing T-DNA. T-DNA is transferred from the bacterium to a host plant which then integrates the genes into its genome. The virulence genes are then expressed which leads to tumour formation. For artificial manipulation, the Ti plasmid is edited so that the opine-synthesis and tumour producing genes are replaced with the genes of interest and selection markers. This edited plasmid is transformed into a strain of *Agrobacterium* that harbours no plasmids prior to infecting the plant cells. The Agrobacterium will then naturally insert the genetic material into the plant cells.[29]

An example of this in modern agriculture is the chimeric crop, *golden rice*.[30] This is genetically modified rice that contains the provitamin A or beta-carotene biosynthetic pathway. The original strain contained three vectors for three genes but now only requires the expression of two genes. The *psy* (phytoene synthase)

gene was derived from daffodil and ctrl (carotene desaturase) from the soil bacterium *Erwinia uredovora*.[30] It was a significant breakthrough in biotechnology and it was the first biosynthetic pathway to be artificially constructed. It also has enormous economic and health impacts, as an estimated 250 million children were affected by vitamin A deficiency in 1995. Beta-carotene derived from golden rice has been proven to be effectively converted to vitamin A in humans and may represent a feasible solution to vitamin A deficiency for developing countries.[31] However, there remains mainstream distrust toward GMOs and as of 2017, golden rice is still not commercially grown as well as being boycotted by environmental activist groups such as Greenpeace.[32] In 2015, as an interesting scientific rebuttal, Kyndt *et al* demonstrated that naturally transgenic accessions of sweet potato have been purposefully selected by humans for millennia. The genome of the modern cultivated sweet potato contains T-DNA sequences. These sequences contain genes that are expressed at detectable levels in tissues. Some of these genes encode for agrocinopine synthase and C-protein which may confer selective advantages. The significance of this discovery is that the cultivated strains of this crop carry particular T-DNA sequences but their wild relatives do not. Kyndt *et al* propose this HGT event conferred some sort of advantage which led to their selection by humans as the modern cultivated strain.[28] This revelation may aid in quelling mainstream perception of GMOs as 'unnatural'. Indeed, as more knowledge of HGT comes to light, we may find that more transgenic crops have been selectively cultivated for millennia due to their positive gene acquisition.

Discussion and Conclusions

Since its discovery in 1928, our understanding of the mechanism of plant HGT has improved greatly but its impact is still undervalued. More verifiable HGT events may now be collected by virtue of a more balanced evaluation method. However, continued breakthroughs are dependent on an expansion of fully sequenced genomes of more neglected species. With a greater collection of sequenced genomes, phylogenetic conflicts may be better identified leading to a more resolved phylogenetic Tree of Life. Unresolved evolutionary events and phylogenetic incongruence in plants may also be resolved through HGT explanations as they have in the past. We may be at the tip of an iceberg in terms of fully grasping the extent of HGT in novel gene acquisition in plants. Mechanisms of plant evolution and speciation that remain cryptic to the scientific community may have answers in HGT as it has been proven before. Further study of HGT may also lead to significant advances in plant biotechnology. New mechanisms of gene transfers may be discovered with further research into host-symbiont relationships. As an unexpected benefit, mainstream hesitation toward transgenic crops may also subside with a growing body of evidence of naturally occurring chimeric crops. HGT clearly deserves more appreciation in the scientific community as its applications in plant biotechnology and implications for evolution are far reaching in impact.

Acknowledgements

This review was partially adapted from a Literature Review I compiled entitled 'Horizontal Gene Transfer to Eukaryotes'. Thank you to Professor Tony Kavanagh for your guidance and advice in this review.

References

1. Leliaert, F., Smith, D.R., Moreau, H., Herron, M.D., Verbruggen, H., Delwiche, C.F. & De Clerck, O. Phylogeny and molecular evolution of the green algae. *Critical Reviews in Plant Sciences.* 2012; 31: 1–46.

2. Griffith, F. The Significance of Pneumococcal Types. *The Journal of Hygiene*. 2012; 27(2), 113–159.

3. Oswald T. Avery, Colin M. MacLeod, Maclyn McCarty. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III. *Journal of Experimental Medicine.* 1994; 79 (2), 137-158;

4. Calum Johnston, Bernard Martin, Gwennaele Fichant, Patrice Polard & Jean-Pierre Claverys. Bacterial Transformation: distribution, shared mechanisms and divergent control. 2014; *Nature Reviews Microbiology,* 12, 181–196.

5. Tomoichiro Akiba, *et al*. On the mechanism of the development of multiple-drug-resistant clones of Shigella. *Japanese Journal of Microbiology.* 1960; 4:219–227.

6. Syvanen, M., & Kado, C. I. Horizontal Gene Transfer. Academic Press. 2002; p. ix.

7. Ragan, M. A., Harlow, T. J., & Beiko, R. G. Do different surrogate methods detect lateral genetic transfer events of different relative ages? *Trends in Microbiology*. 2006; 14(1), 4-8.

8. Worning, P. Structural analysis of DNA sequence: evidence for lateral gene transfer in Thermotoga maritima. *Nucleic Acids Research.* 2002; 28(3), 706-709.

9. Lawrence, J. G., & Ochman, H. Reconciling the many faces of lateral gene transfer. Trends in Microbiology. 2002; 10(1), 1-4.

10. Lawrence, J. G., & Ochman, H. Amelioration of Bacterial Genomes: Rates of Change and Exchange. *Journal of Molecular Evolution.* 1997; 44(4), 383-397.

11. Deschavanne, P., & Filipski, J. Correlation of GC content with replication timing and repair mechanisms in weakly expressed E.coli genes. *Nucleic Acids Research.* 1995; 23(8), 1350-1353.

12. Wuitschick, J. D., & Karrer, K. M. Analysis of Genomic G C Content, Codon Usage, Initiator Codon Context and Translation Termination Sites In Tetrahymena Thermophila. *The Journal of Eukaryotic Microbiology.* 1999; 46(3), 239-247

13. Guindon, S., & Perriere, G. Intragenomic Base Content Variation Is a Potential Source of Biases When Searching for Horizontally Transferred Genes. *Molecular Biology and Evolution.* 2001; 18(9), 1838-1840.

14. Keeling, P. J., & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics*. 2008; 9(8), 605-618.

15. Andersson, J. O. Eukaryotic Gene Transfer: Adaptation and Replacements. Horizontal Gene Transfer in the Evolution of Pathogenesis. *Cambridge University Press.* 2008; 293-316.

16. Richardson, A. O., & Palmer, J. D. Horizontal gene transfer in plants. *Journal of Experimental Botany.* 2006; 58(1), 1-9.

17. Huang, J. Horizontal gene transfer in

eukaryotes: The weak-link model. *BioEssays*. 2017.

18. Lacroix, B., & Citovsky, V. Transfer of DNA from Bacteria to Eukaryotes. *MBio.* 2016; 7(4).

19. Zimorski, V., Ku, C., Martin, W. F., & Gould, S. B. Endosymbiotic theory for organelle origins. *Current Opinion in Microbiology*. 2014; 22, 38-48.

20. Bock, R. The give-and-take of DNA: horizontal gene transfer in plants. *Trends in Plant Science*. 2010; 15(1), 11-22.

21. Sanchez-Puerta, M. V., Cho, Y., Mower, J. P., Alverson, A. J., & Palmer, J. D. Frequent, Phylogenetically Local Horizontal Transfer of the cox1 Group I Intron in Flowering Plant Mitochondria. *Molecular Biology and Evolution.* 2008; 25(8), 1762-1777.

22. Rice, D. W., Alverson, A. J., Richardson, A. O., Young, G. J., Sanchez-Puerta, M. V., Munzinger, J., . . . Palmer, J. D. Horizontal Transfer of Entire Genomes *via* Mitochondrial Fusion in the Angiosperm Amborella. *Science.* 2013; 342(6165), 1468-1473.

23. Singer, K., Shiboleth, Y. M., Li, J., & Tzfira, T. Formation of Complex Extrachromosomal T-DNA Structures in Agrobacterium tumefaciens-Infected Plants. *Plant Physiology.* 2012; 160(1), 511-522.

24. Liang, Z., & Tzfira, T. In vivo formation of double-stranded T-DNA molecules by T-strand priming. *Nature Communications.* 2013; 4.

25. Pitzschke, A., & Hirt, H. New insights into an old story: Agrobacterium-induced tumour formation in plants by plant transformation. *The EMBO Journal*. 2010; 29(6), 1021-1032.

26. Lacroix, B., & Citovsky, V. A Functional Bacterium-to-Plant DNA Transfer Machinery of Rhizobium etli. *PLOS* Pathogens. 2016; 12(3).

27. Suzuki, K., Yamashita, I., & Tanaka, N. Tobacco plants were transformed by Agrobacterium rhizogenes infection during their evolution. *Plant J.* 2002; 32(5):775-87.

28. Kyndt, T., Quispe, D., Zhai, H., Jarret, R., Ghislain, M., Liu, Q., . . . Kreuze, J. F. The genome of cultivated sweet potato contains Agrobacterium T-DNAs with expressed genes: An example of a naturally transgenic food crop. *Proceedings of the National Academy of Sciences*. 2015; 112(18), 5844-5849.

29. Gelvin, S. B. Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. *Microbiology and Molecular Biology Reviews*. 2003; 67(1), 16-37.

30. Ye, X. . Engineering the Provitamin A (Beta-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science.* 2000; 287(5451), 303-305.

31. Tang, G., Qin, J., Dolnikowski, G. G., Russell, R. M., & Grusak, M. A. Golden Rice is an effective source of vitamin A. *American Journal of Clinical Nutrition.* 2009; 89(6), 1776-1783.

32. Greenpeace. Special Report: Golden Rice [Internet]. . Available from: http://www.greenpeace.org/international/en/campaigns/agriculture/problem/Greenpeace-and-Golden-Rice/

Mechanisms of Immune Tolerance in the Development of Successful Pregnancy

Stephen Cunningham
Junior Sophister
Immunology

In the context of immunology, pregnancy presents a paradoxical concept. Under normal conditions, the host immune system recognises foreign antigens from pathogens and transformed cells. Part of the immune response to these antigens involves their presentation to T cells through cell surface machinery known as the Major Histocompatibility Complex (MHC). Recognition of nonself-MHC, that is, MHC derived from another individual, is what dictates transplant success – if the MHC of the graft is not identical to that of the host, it will be recognised as foreign tissue and undergo immune rejection. An allograft is one which is transplanted from a genetically non-identical donor of the same species. It follows that a foetus in utero may be considered an allograft, due to 50% of its genes being paternally-derived and, as such, it is expected that the maternal immune system would reject it. However, we know that this is not the case, as most pregnancies are brought to term, implying that there a number of immune tolerance mechanisms employed by the mother and foetus to maintain an immunologically-privileged site. If compromised, the maternal immune system recognises the allogeneic foetal antigens as foreign and mounts a response, potentially resulting in severe complications such as miscarriage, premature delivery and preeclampsia. Miscarriage occurs in about 15% of pregnancies, and remains recurrent in 2-5% of cases, many of which go unexplained. It is thought that if these have an immunological basis, and if the mechanisms of immune tolerance can be elucidated, potential therapies may follow. The purposes of this review is to discuss some of these tolerance mechanisms, including suppression of maternal alloreactivity by placental phosphocholinated peptides, myeloid-derived suppressor cells (MDSCs), remodelling of decidual vasculature by uterine natural killer (uNK) cells, upregulation of maternal regulatory T cells (Tregs), as well as the consequences of a breakdown in immune tolerance during pregnancy.

Phosphocholinated Peptides

Phosphocholine is a precursor of phosphatidylcholine, an essential phospholipid component of cell membranes. It has been shown that filarial nematodes, parasites that cause persistent infection in approximately 100 million people in the tropics,[1] employ phosphocholine-linked glycoproteins as an immune evasion mechanism. This modulation of the immune response by excretory-secretory (ES) phosphocholinated peptides towards an anti-inflammatory phenotype affects components of both the innate and particularly the adaptive immune system; including T cells, B cells, dendritic cells, mast cells, macrophages and complement activation.[23] The mechanism whereby the phosphocholinated peptide ES-62, a major excretory-secretory product of A*canthocheilonema viteae*,[4] suppresses B and T cell function is by essecntially desensitising the cells to activation. This results in uncoupling of the B cell receptor (BCR) and T cell receptor (TCR) from proliferative transduction pathways by providing a signal insufficient to cause proliferation but adequate to anergise them.[1] Such pathways include phosphoinositide 3-kinase (PI3K)- and Ras-mitogen-activated protein kinase (MAPK)-dependent transduction in B cells, and phospholipase D-, protein kinase C-, PI3K- and Ras-MAPK-dependent cascades in T cells.[5]

The immunomodulatory function of phosphocholinated peptides is dependent on the phosphocholine group, as phosphocholine-linked bovine serum albumin (BSA) and chloride mimics the action of ES-62.

Research by Lovell *et al*. in 2007[6] identified novel placental secretory polypeptides containing phosphocholine as a post-translational modification. Neurokinin B and corticotrophin-releasing factor (CRF), as well as precursors of hemokinin, activin, adrenocorticotrophin and follistatin have all been identified as placental secretory phosphocholinated peptides. It follows that, since the presence of phosphocholine groups on nematode ES peptides suppresses the host immune response to the parasite, perhaps phosphocholinated peptides of the placenta serves a similar function with the foetus. Because the placenta is derived from embryonic trophoblast cells, it, like the foetus, is also considered semiallogeneic.

CTP:PC cytidylyltransferase-α is an enzyme expressed ubiquitously throughout the body, and is responsible for the introduction of phosphocholine into phosphatidylcholine during membrane formation.[6] However, in 1998, Lykidis *et al*.[7] discovered a second isoform of the enzyme; CTP:PC cytidylyltransferase-β. Furthermore, it was found that this enzyme is expressed in large amounts in the placenta and testis relative to other tissues.[8] As both the placenta and testis represent immunologically-privileged sites, it is thought that this enzyme is responsible for post-translational phosphocholine modifications of secretory peptides that ultimately suppress the maternal immune system and protect the foetus from immune rejection.[6]

As stated previously, phosphocholinated nematode ES-62 desensitises B cells to activation. At high concentrations, however, ES-62 is mitogenic for B cells *in vitro*.[4] As pregnancy approaches term in humans, expression of placental phosphocholinated peptides in maternal circulation increases, perhaps having a similar effect on B cell proliferation. An increase in B cell activity in the mother is essential, as this increases production of immunoglobulin-G (IgG), an antibody which can pass through the placenta to provide passive immunity to the neonate *in utero*.[9]

Myeloid-derived suppressor cells (MDSCs)

Haematopoiesis is the process whereby blood cells differentiate from haematopoietic stem cells (HSCs) in the bone marrow. Immune cells are generated from one of two pathways – myeloid progenitor cells generate macrophages, dendritic cells, monocytes, granulocytes and mast cells, and lymphoid progenitors generate T cells, B cells and natural killer cells.

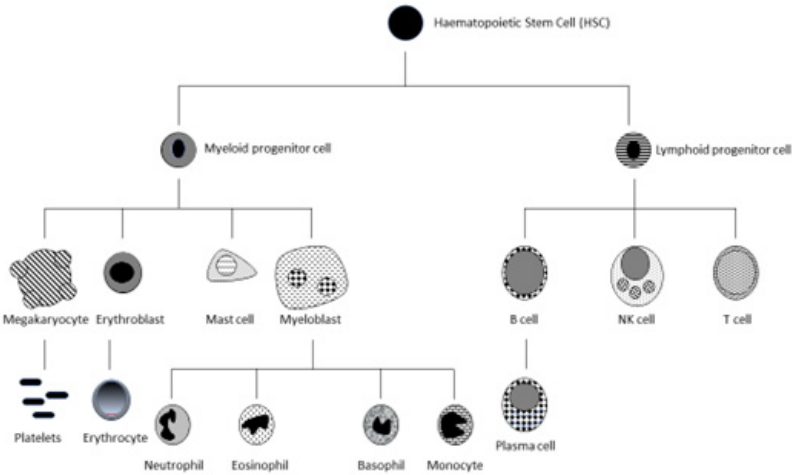


Figure 1. Shows the haematopoietic pathway, whereby cells of the blood are differentiated from haematopoietic stem cells in the bone marrow. Cells can follow one of two pathways; myelopoiesis, which generates the majority of innate immune cells, or lymphopoiesis, which generates NK cells and cells of the adaptive immune system.
Adapted from Doulatov et al. (2012)[10]

Myeloid-derived suppressor cells (MDSCs) are a heterogenous subset of monocytic or granulocytic cells that locally suppress the immune system.[11] Granulocytic MDSCs are expressed in the placenta and umbilical cord, where they suppress T cell cytotoxicity, induce Treg cell proliferation and inhibit NK cell activation *via* expression of reactive oxygen species, arginase-1 and nitric oxide synthase.[11,12].

A study by Kang *et al*. in 2016 analysed MDSC levels in the decidua of humans undergoing elective abortion compared to spontaneous abortion (i.e. miscarriage). They found that granulocytic MDSC (G-MDSC) levels were significantly lower in patients who had miscarried relative to those undergoing elective abortion, but monocytic MDSC (M-MDSC) levels were unchanged. They also examined mouse uteri, and found that non-pregnant mice expressed little to no G-MDSCs, but populations rapidly expanded during pregnancy.[13]

Another study carried out by Ostrand-Rosenberg *et al*. (2016)[14] examined the effect of depleting MDSCs on implantation of mice embryos. They found that, depletion of cells with an MDSC phenotype resulted in implantation failure and increased T cell activity in the uterus. Interestingly, upon restoration of MDSCs, they observed restoration of successful pregnancy, downregulation of L-selectin expression on naïve T cell surfaces, reducing their ability to enter lymph nodes and become activated.[14]

It is thought that MDSCs may occupy a central role in the area of reproductive immunology. It is unclear how approximately 50% of miscarriages occur, and if MDSCs are found to play a role in these cases, they may provide a therapeutic target in the treatment of women suffering from recurrent miscarriages.

Uterine Natural Killer (uNK) Cells

Natural killer cells are cytotoxic lymphocytes which play a critical role in the innate immune response to viruses and transformed cells. They are large granular cells that constitute the third type of cells differentiated from the common lymphoid progenitor, along with B and T cells. NK cells account for about 10-15% of peripheral blood mononuclear cells.[15] Their primary effector function in the immune system is cytotoxicity, however, they are also important immunoregulatory cells, as can be seen from their role in the development of successful pregnancy in eutherian mammals.

Uterine natural killer (uNK) cells are a subset of NK cells found in the uterus with a $CD56^{bright}$+$CD16^{dim}$+ surface antigen phenotype.[16] They have a similar morphology to peripheral NK cells, with characteristic granules containing perforin and granzymes needed for cytotoxicity, but exhibit a much lower level of cytotoxicity; attributable to their inability to polarise their microtubule-organising centres and perforin-containing granules.[17] uNK cells are present in the non-pregnant endometrium and decidua, and at a much lower level in the endometrium of premenarcheal and post-menopausal women; implying that their primary effector function is required by fertile women.

It has also been shown that approximately 40% of uNK cells express Ki67, a cellular marker of proliferation, indicating that these cell populations expand locally in the uterus. Interestingly, uNK cell numbers are proportional to progesterone levels, a

hormone which induces upregulation of IL-15 in these cells.[18] IL-15 has been shown to stimulate proliferation of uNK cells in vitro, further supporting the idea that these cells proliferate and expand their population locally rather than migrating to the uterus from the periphery. Approximately 70% of decidual lymphocytes are CD56+, primarily found around maternal spiral arteries, reflecting their fundamental importance in the uterus.[19]

The role of uNK cells in the establishment of successful pregnancy is believed to be *via* co-operation with extravillous trophoblast (EVT) cells of the foetus, which are responsible for penetrating the media surrounding the uterine spiral arteries to ensure a sufficient and appropriate blood flow is supplied to the developing foetus.[20] Flaws in this process is what contributes to the great obstetric syndromes (GOS), namely miscarriage, preeclampsia and foetal growth restriction.[21] In mice, this process is mediated by secretion of IFN-γ by uNK cells,[22], but it is much more complex in humans. Remodelling of decidual vasculature and subsequent placentation in humans is thought to be dependent on recognition of foetal EVT cells by uNK cells. This recognition facilitates secretion of angiogenic growth factors by uNK cells, which subsequently mediate redirection of the bloodflow to the foetus.[23] EVTs extend from the foetus and penetrate the decidua, coming into close proximity with uNK cells and subsequently become susceptible to being recognised as allografts. This allogeneic recognition may, ironically, be the reason why decidual vasculature remodelling is successful. Foetal EVTs secrete soluble human leukocyte antigens (HLA)-C, -E and –G, which are thought to ligate with uNK killer cell Ig-like receptors (KIRs), triggering the production of aforementioned angiogenic growth factors.[16]

HLA-G, for example, binds to KIR2DL4, a uNK cell receptor which induces signal transduction pathways that direct the cell towards an angiogenic phenotype.[24] Such growth factors include angiopoietin-1, angiopoietin-2, TGF-β1 and vascular endothelial growth factor (VEGF)-C.[23]

Facilitated by these pro-angiogenic factors, EVT cells cause fibrinoid change (remodelling of connective tissue) by destruction of the arterial muscle, inhibiting their ability to vasocontrict, followed by foetus-derived endovascular trophoblast cells taking the place of the destroyed endothelium.[20]

Overall, it is clear that uNK cells play an essential role in the initiation of successful placentation by mediating remodelling of uterine spiral arteries by EVT cells. As mentioned previously, this process is what contributes to the development of the GOS, perhaps indicating that uNK cells may provide a potential therapeutic target in the treatment of recurrent miscarriage, preeclampsia and foetal growth restriction.

Regulatory T (Treg) Cells

Treg cells are an important heterogenous subset of T lymphocytes, discovered by Sakaguchi *et al*. in 1995.[25] They have an important role in regulating inflammatory immune responses, by suppressing activation and effector functions of other immune cells, particularly T cells, preventing autoimmunity, chronic inflammation and maintaining homeostasis.[26] In this way, Treg cells are responsible for limiting the intensity and duration of a T cell response to both alloantigens and exogenously-derived antigens; and their increased populations at the maternal-foetal interface during pregnancy provides a mechanism whereby the developing foetus can be protected from immune rejection.[27]

Although Treg cells primarily regulate CD4+ and CD8+ T cell responses, they are also involved in the suppression of B cell proliferation and activation, NK cell cytotoxicity, and macrophage and dendritic cell maturation.[27] Through this regulation of both the innate and adaptive immune system, Treg cells protect the foetus from a wide range of potentially fatal reactions.

Treg cells are identifiable by their expression of the transcription factor FOXP3, as well as their CD4+CD25+ surface antigen phenotype. Their effector functions are mediated by secretion of IL-10 and TGFβ, two anti-inflammatory cytokines involved in tissue remodelling and the resolution of the immune response. They also act as a sink for IL-2, an essential cytokine required for T cell proliferation.[27]

It has been shown that maternal human chorionic gonadiotropin recruits Treg cells to the decidua, and foetal trophoblast-derived CCL5 enhances expression of FOXP3.[28,29] Furthermore, Prieto *et al*. (2006)[30] showed that estradiol, a hormone expressed at elevated levels during pregnancy, enhances the suppressive function of Treg cells in utero, amplifying local immune tolerance between the mother and foetus.

As uNK cells, Treg cells also express Ki67, a marker of proliferation, indicating that these populations expand locally rather than being recruited from the periphery.[26]

CD4+CD25+FOXP3+ Treg cell populations are reduced in both the circulation and decidua in women suffering from recurrent miscarriages and preeclampsia.[26] A reduction in Treg cells facilitates an increase in the activity of Th17 cells, which trigger an inflammatory immune response, possibly initiating foetus rejection.[31] It follows that, by analysis of CD25+FOXP3+ cell markers, it may be possible to predict miscarriages in women with a history of failed pregnancies.

The role of Treg cells in the development of normal pregnancy is similar to that of uNK cells. Their regulatory function provides local immune suppression at the maternal-foetal interface, inhibiting allorecognition of the foetus and maintaining an immunologically-privileged site in healthy pregnancies.

Conclusions

Overall, it is evident that there are a number of mechanisms employed by both the mother and foetus that create immune tolerance between them, allowing for the development of normal pregnancy. A breakdown in this immune-privilege, as discussed, can lead to a variety of potentially fatal outcomes, such as miscarriage, preeclampsia and premature delivery. A chronic or sustained absence in any of these local immunosuppressive mechanisms can cause recurrent disorders in successive pregnancies, such as recurrent miscarriages and even infertility.

Phosphocholinated peptides, myeloid-derived suppressor cells, uterine natural killer cells and regulatory T cells provide different methods of immune suppression with similar outcomes, whereby the uterus is maintained as an immunologically-privileged site. Functionally, they act in synergy to create cumulative local immune modulation; preventing allogeneic-mediated immune rejection of the foetus and placenta.

T cells, B cells, dendritic cells, mast cells, macrophages, dendritic cells and the complement system are regulated by Treg cells and phosphocholinated peptides such as neurokinin B, corticotrophin-releasing factor (CRF), and precursors of hemokinin, activin, adrenocorticotrophin and follistatin. MDSCs modulate T cell cytotoxicity and inhibit NK cell activation *via* expression of reactive oxygen species, arginase-1 and nitric oxide synthase, and also induce Treg cell proliferation. From this, it is clear that inflammatory immune responses are cumulatively regulated by phosphocholinated peptides as well as MDSCs and Treg cells; the latter of which is induced to expand by the former. These immunomodulatory mediators, in tandem with uNK cells, represent a complex network that maintain the delicate balance between protecting the neonate from the maternal immune system, yet still facilitating passive immunity by allowing IgG to pass the placental syncytium barrier. Maintenance of this balance requires a dynamic system of regulation, mirroring that of our immune systems when it comes to mounting an appropriate immune response to a stimulus with subsequent resolution to prevent autoimmunity and chronic inflammatory diseases.

In conclusion, the mechanisms described in this review represent potential targets which, if further elucidated, may provide therapies in the treatment of pregnancy-related pathologies.

Acknowledgements

The author would like to thank Mr. J. Schofield, Ms. K. Hughes and Mr. C. McLoughlin for helpful edits and advice.

References

1. Harnett W. A phosphorylcholine-containing filarial nematode secreted product disrupts lymphocyte activation by targetting key proliferative signalling pathways. *Immunology Letters.* 1997;56(1-3):355-356

2. Al-Riyami LHarnett W. Immunomodulatory Properties of ES-62, a Phosphorylcholine - Containing Glycoprotein Secreted by Acanthocheilonema viteae. Endocrine, Metabolic & Immune Disorders-*Drug Targets.* 2012;12(1):45-52.
3. Ahmed U, Maller N, Iqbal A, Al-Riyami L, Harnett W, Raynes J. The Carbohydrate-linked Phosphorylcholine of the Parasitic Nematode Product ES-62 Modulates Complement Activation. *Journal of Biological Chemistry.* 2016;291(22):11939-11953.

4. Harnett W, Houston K, Amess R, Worms M. Acanthocheilonema viteae: Phosphorylcholine Is Attached to the Major Excretory-Secretory Product *via* an N-Linked Glycan. *Experimental Parasitology.* 1993;77(4):498-502.

5. Harnett W, Harnett M. Modulation of the host immune system by phosphorylcholine-containing glycoproteins secreted by parasitic filarial nematodes. Biochimica et Biophysica Acta (BBA) - *Molecular Cell Research*. 2001;1539(1-2):7-15.

6. Lovell T, Woods R, Butlin D, Brayley K, Manyonda I, Jarvis J. Identification of a novel mammalian post-translational modification, phosphocholine, on placental secretory polypeptides. *Journal of Molecular Endocrinology.* 2007;39(3):189-198.

7. Lykidis A, Murti K, Jackowski S. Cloning and Characterization of a Second Human CTP:Phosphocholine Cytidylyltransferase. *Journal of Biological Chemistry.* 1998;273(22):14022-14029.

8. Lykidis A, Baburina I, Jackowski S. Distribution of CTP:Phosphocholine Cytidylyltransferase (CCT) Isoforms: IDENTIFICATION OF A NEW CCT SPLICE VARIANT. *Journal of Biological Chemistry.* 1999;274(38):26992-27001.

9. Kane Sacquah L. Placental Transport of Immunoglobulins: A Clinical Review for Gastroenterologists Who Prescribe Therapeutic Monoclonal Antibodies to Women During Conception and Pregnancy. *The American Journal of Gastroenterology.* 2009;104(1):228-233.

10. Doulatov S, Notta F, Laurenti E, Dick J. Hematopoiesis: A Human Perspective. *Cell Stem Cell.* 2012;10(2):120-136.

11. Gantt S, Gervassi A, Jaspan H, Horton H. The Role of Myeloid-Derived Suppressor Cells in Immune Ontogeny. *Frontiers in Immunology.* 2014;5.

12. Pan T, Liu Y, Zhong L, Shi M, Duan X, Wu K *et al*. Myeloid-derived suppressor cells are essential for maintaining feto-maternal immunotolerance *via* STAT3 signaling in mice. *Journal of Leukocyte Biology.* 2016;100(3):499-511.

13. Kang X, Zhang X, Liu Z, Xu H, Wang T, He L *et al*. Granulocytic myeloid-derived suppressor cells maintain feto-maternal tolerance by inducing Foxp3 expression in CD4+CD25−T cells by activation of the TGF-β/β-catenin pathway. *Molecular Human Reproduction*. 2016;22(7):499-511.

14. Ostrand-Rosenberg S, Sinha P, Figley C, Long R, Park D, Carter D *et al*. Frontline Science: Myeloid-derived suppressor cells (MDSCs) facilitate maternal–fetal tolerance in mice. *Journal of Leukocyte Biology.* 2016;:jlb.1HI1016-306RR.

15. Gras Navarro A, Björklund A, Chekenya M. Therapeutic Potential and Challenges of Natural Killer Cells in Treatment of Solid Tumors. *Frontiers in Immunology.* 2015;6.

16. Zhang J, Chen Z, Smith G, Croy B. Natural killer cell-triggered vascular transformation: maternal care before birth?. *Cellular and Molecular Immunology.* 2010;8(1):1-11.

17. Kopcow H, Allan D, Chen X, Rybalov B, Andzelm M, Ge B *et al*. Human decidual NK cells form immature activating

synapses and are not cytotoxic. *Proceedings of the National Academy of Sciences.* 2005;102(43):15563-15568.

18. Croy B. Uterine natural killer cells: insights into their cellular and molecular biology from mouse modelling. *Reproduction*. 2003;126(2):149-160.

19. Whitley G, Cartwright J. Cellular and Molecular Regulation of Spiral Artery Remodelling: Lessons from the Cardiovascular Field. *Placenta*. 2010;31(6):465-474.

20. Moffett A, Colucci F. Uterine NK cells: active regulators at the maternal-fetal interface. *Journal of Clinical Investigation.* 2014;124(5):1872-1879.

21. Brosens I, Pijnenborg R, Vercruysse L, Romero R. The "Great Obstetrical Syndromes" are associated with disorders of deep placentation. *American Journal of Obstetrics and Gynecology*. 2011;204(3):193-201.

22. Redline R. Role of Uterine Natural Killer Cells and Interferon γ in Placental Development. *The Journal of Experimental Medicine.* 2000;192(2):F1-F4.

23. Lash G. Expression of angiogenic growth factors by uterine natural killer cells during early pregnancy. *Journal of Leukocyte Biology.* 2006;80(3):572-580.

24. Rajagopalan S, Long E. KIR2DL4 (CD158d): An activation receptor for HLA-G. *Frontiers in Immunology.* 2012;3.

25. Takahashi T, Sakaguchi S. Naturally Arising CD25+CD4+ Regulatory T Cells in Maintaining Immunologic Self-Tolerance and Preventing Autoimmune Disease. *Current Molecular Medicine.* 2003;3(8):693-706

26. Ernerudh J, Berg G, Mjösberg J. Regulatory T Helper Cells in Pregnancy and their Roles in Systemic versus Local Immune Tolerance. *American Journal of Reproductive Immunology*. 2011;66:31-43.

27. Guerin L, Prins J, Robertson S. Regulatory T-cells and immune tolerance in pregnancy: a new target for infertility treatment?. *Human Reproduction Update.* 2009;15(5):517-535.

28. Fraccaroli L, Alfieri J, Larocca L, Calafat M, Mor G, Leiros C. A potential tolerogenic immune mechanism in a trophoblast cell line through the activation of chemokine-induced T cell death and regulatory T cell modulation. *Human Reproduction.* 2008;24(1):166-175.

29. Schumacher A, Brachwitz N, Sohr S, Engeland K, Langwisch S, Dolaptchieva M. Human Chorionic Gonadotropin Attracts Regulatory T Cells into the Fetal-Maternal Interface during Early Human Pregnancy. *The Journal of Immunology.* 2009;182(9):5488-5497.

30. Prieto G, Rosenstein Y. Oestradiol potentiates the suppressive function of human CD4+ CD25+ regulatory T cells by promoting their proliferation. *Immunology.* 2006;118(1):58-65.

31. Winger E, Reed J. Low circulating CD4+CD25+Foxp3+ T regulatory cell levels predict miscarriage risk in newly pregnant women with a history of failure. *Journal of Reproductive Immunology.* 2010;86(2):110.

TS
SR

The Regulation of glmS Expression as an Example of the Control of mRNA Stability in Bacteria

Michaela Kearney
Senior Sophister
Genetics

Prokaryotic regulation of gene expression, and thus, the regulation of protein levels in the cell, are commonly associated with the control of a promoter element during the transcription stage of gene expression. Recently, however, much research has focused on post-transcriptional regulation. This typically involves control of the stability of mRNA in the cell, which may or may not lead to its degradation. This provides the cell with extra control of gene expression, thus enabling it to fine-tune its protein levels in order to meet its needs in specific environments. This review provides an overview of the mechanism and components involved in mRNA degradation, and focuses on the expression of glmS to provide an insight into the diverse mechanisms and roles of regulating mRNA stability in control of gene expression.

Introduction

Transcriptional regulation of gene expression is an area that has been extensively studied in both prokaryotes and eukaryotes. It was long assumed to be the primary mechanism of regulating gene expression to meet the cell's needs. However, disagreements between the amount of mRNA in a cell and the concentration of its corresponding protein suggest a strong role of post-transcriptional regulation in gene expression. This adds another layer of control which allows the cell to fine-tune its gene expression in order to adapt to its surroundings. In addition, mRNA degradation, and thus translational regulation, is a rapid process This is particularly advantageous in stressful conditions, such as extreme temperature or pressure, where the cell must rapidly produce the proteins required for survival.

The regulation of stability is generally achieved through a combination of the action of *cis*-encoded features of the mRNA transcript sequence and interaction with ribonucleases to stimulate or inhibit degradation. The efficiency of translation into protein also contributes to mRNA turnover, and this can be regulated by influencing ribosome-binding to the ribosome binding site (RBS). A wide range of components are involved in mRNA turnover, including RNases, small non-coding RNAs (sRNAs) and a variety of protein cofactors or chaperones, such as the chaperone Hfq.[1,2,3] These components cooperate as part of an extensive network to ensure efficient mRNA stabilisation or degradation.

Many mRNA regulatory pathways involve multiple sRNAs and proteins, often working hierarchically. Other mechanisms can be considered to be more simplistic, sometimes involving just the mRNA transcript itself, without the use of additional proteins. This process, where structures and motifs within the transcript are capable of mediating regulation of the transcript itself, is known as cis-encoded regulation. The regulation of *glmS* provides an example of both mechanisms. A small number of sRNAs are utilised in a hierarchial manner in Gram-negative bacteria, while *cis*-encoded regulation is employed by Gram-positive bacteria.

The purpose of this review is to provide an insight into the regulation of mRNA stability in the control of prokaryotic gene expression. The mechanism and components involved in mRNA degradation will be briefly described, followed by an overview of the regulation of the expression of the *glmS* gene. Its regulation in both Gram-positive and Gram-negative bacteria will be discussed in order to highlight the versatility of the mRNA regulatory machinery.

mRNA Degradation: Mechanism and Machinery

The half-life of prokaryotic mRNA is considerably shorter than that of eukaryotes, with most mRNAs in *E. coli* having a half-life of approximately two minutes at 37°C.[4] A number of key factors are involved in mRNA degradation, which regularly cooperate to ensure efficient gene regulation. The major enzymes involved are endoribonucleases, which cleave the transcript at an internal site, and exoribonucleases, which remove nucleotides from either end of a transcript. One of the most prominent endoribonuclease in *E. coli* is RNase E. RNase E, along with enzymes such as RNA helicase B (RhlB), enolase and the exoribonuclease polynucleotide phosphorylase (PNPase), forms the degradosome machinery, which is responsible for mRNA decay (Figure 1).[5]

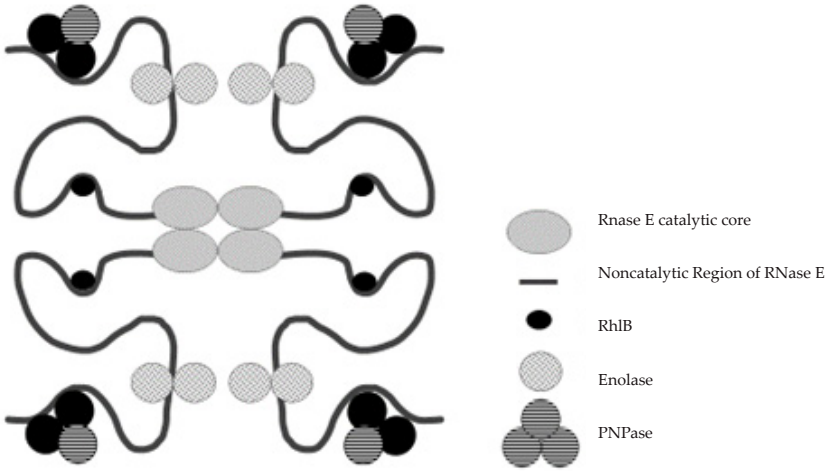


Figure 1: The RNase E degradosome. Simplified diagram of the structure and components of the RNase E degradosome. The subunits of the degradasome are indicated The noncatalytic region has been drawn symmetrically, however it is a dynamic structure. Thus, each extension would form a random coil and would act independently of each other
Figure adapted from Carpousis, 2007.[5]

The degradation procedure in *E. coli* is thought to initiate with internal cleavage mediated by the RNase E degradosome. RNase E is a 5′-dependent endoribonuclease, with a preference for a mono-phosphorylated 5′ end over a tri-phosphorylated end.[6] Therefore, it is thought that, similar to the decapping of eukaryotic mRNAs, a pyrophosphatase converts the tri-phosphate cap to a mono-phosphate cap before endonucleolytic cleavage.[7] Before cleavage the transcript typically contains a hairpin structure at its 3′ end preventing degradation. This structure is formed when 2 regions of the same strand are complementary to each other when read in opposite directions, resulting in the formation of a double helix ending in an unpaired loop. Following cleavage, however, the 3′ ends of the generated strands are unprotected, allowing further degradation by exoribonucleases, RhlB and poly(A) polymerase (PAP).[8,9] PAP adds poly(A) tails to the 3′ ends of mRNAs, marking them for degradation. The mechanism by which PAP targets transcripts is not fully understood. It is suggested that sequence recognition does not play a significant role in PAP recruitment, however certain structural features of the 5′ end, such as mono-phosphorylation, have been shown to positively influence the efficiency of PAP-mediated polyadenylation.[10,11,12] This is an example of selective processing of mRNAs in order to express the genes required in specific environments.

Other factors involved in mRNA degradation are small RNAs, which typically operate by binding to their target mRNAs on or near the Shine-Dalgarno sequence, a specific sequence of nucleotides which is responsible for ribosome recruitment.

Binding of sRNAs to this region blocks access of the ribosome, thereby hindering translation.[2] The role of sRNAs in regulation of gene expression is discussed later in this review.

The Regulation of Expression of *glmS*

The *glmS* gene encodes glucosamine-6-phosphate (GlcN-6-P) synthase, an enzyme involved in amino-sugar metabolism. It is found in many bacterial species and is post-transcriptionally regulated in response to GlcN-6-P, the substrate of the enzyme. The mechanisms by which it is regulated differ in Gram-positive and Gram-negative bacteria, and so provide prime examples of the variety of ways in which mRNA stability can be regulated.

Gram-Positive Bacteria

In Gram-positive bacteria, the mRNA transcript contains a ribozyme at the 5′ end. In the presence of GlcN-6-P, the ribozyme undergoes self-cleavage, inactivating the transcript and preventing gene expression.[13,14] This is an example of *cis*-encoded regulation and is one of the most well-studied post-transcriptional regulation systems in prokaryotes. The use of a ribonuclease, RNase J1, is of note in this system. Until the discovery of RNase J1, it was assumed that bacterial degradation could only occur in a 3′ to 5′ direction. However, RNase J1 proceeds in a 5′ to 3′ direction.[15] The *glmS* gene of *B. subtilis* falls into the category of riboswitch genes, that is, genes which contain a metabolite-binding domain that induces a conformational change upon binding.[16] Thus, this is an example of post-transcriptional gene regulation that does not require the assistance of sRNAs. The metabolite in question is GlcN-6-P, the substrate of GlcN-6-P synthase. In an experiment designed to investigate the riboswitch activity of *glmS* RNA, Winkler *et al*. used an in-line probing assay, which is based on the principle that unstructured regions of RNA are less stable than regions involved in secondary and tertiary structures, and thus will undergo degradation at a faster rate. The results showed that the *glmS* transcript underwent cleavage at a much higher rate than other known ribozymes, suggesting the simultaneous use of multiple catalytic strategies to accelerate the self-cleavage process. By incubating *glmS* RNA with various metabolites and comparing the cleavage rates, the researchers were able to confirm ribozymatic activity of the transcript. Furthermore, it was shown to have a specificity for GlcN-6-P. Following on from this, various mutations were introduced into the *glmS* RNA, and their resulting ribozymatic activities *in vitro* were compared with their ability to control gene expression *in vivo*, using a β-galactosidase reporter gene. The resulting negative correlation confirmed the role of the ribozymatic activity of *glmS* in its own repression.[13] A simplistic overview of this mechanism is given in Figure 2.

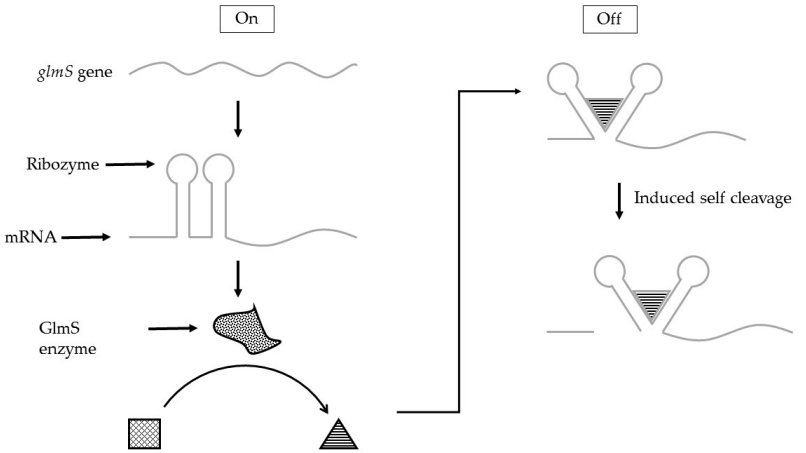


Figure 2: Riboswitch activity of *glmS*. The *glmS* gene is shown as an orange curved line. It is transcribed into an mRNA transcript with ribozymatic activity. Translation of the mRNA produces the *GlmS* enzyme, which catalyses the synthesis of GlcN-6-P. GlcN-6-P then binds to the *glmS* transcript, inducing a conformational change and consequently self-cleavage of the transcript, thus inactivating it
Figure redrawn from Cech, 2004

Further studies investigated the role of RNase J1 by analysing ribozyme-containing transcripts in the presence and absence of the ribonuclease. The level of ribozyme-containing transcripts was found to increase upon removal of RNase J1, highlighting its importance in the degradation of *glmS* mRNA. The results supported the idea that RNase J1 is responsible for degradation of the cleaved *glmS* mRNA following self-cleavage. Furthermore, a *glmS* variant, known as M12, was generated which was identical to the wild-type 3′ cleavage product except for the presence of a triphosphate group at its 5′ terminus, while wild-type cleavage products display a 5′ hydroxyl group. The levels of M12 accumulated in the presence and absence of RNase J1, in contrast to the wild-type cleavage products which only accumulated in the absence of RNase J1. This suggested the recognition of 5′ hydroxyl structures by RNase J1. To confirm this, a hammerhead ribozyme-*rpsT* fusion was inserted into *E. coli*, followed by expression of RNase J1. Similar to the *glmS* ribozyme, the hammerhead ribozyme also generates products with a 5′ hydroxyl terminus. Following cleavage, RNase J1 efficiently destabilises the 3′ cleavage product, confirming the ability of the ribonuclease to recognise and cleave substrates with a 5′ hydroxyl group.[17] Thus, the *glmS* system in *B. subtilis* is a prime example of regulation of mRNA decay in the control of gene expression.

Gram-Negative Bacteria

In Gram-negative bacteria such as *E. coli*, there is no ribozyme in the *glmS* transcript. Thus, regulation occurs *via* a different mechanism, namely a *trans*-encoded mechanism. In *E. coli*, *glmS* is transcribed as part of the *glmUS* operon, which is then cleaved by ribonuclease E to generate a monocistronic *glmS* mRNA. GlmU encodes UDP-GlcNAc synthetase, a key component in cell wall synthesis. This protein is constantly needed in the cell, and thus *glmU* is constitutively expressed. The expression of *glmS*, in contrast, is only needed when external amino sugars are in short supply.[18] As *glmS* and *glmU* are coexpressed, it follows that the expression of *glmS* must be regulated at a post-transcriptional level. This is indeed the case: The sRNAs GlmY and GlmZ are positive regulators of the *glmS* mRNA.[19, 20] In addition, two proteins, PAP I and YhbJ (encoded by *pcnB* and *yhbJ*, respectively), are known to be negative regulators of *glmS* expression.[19,21]

Through a series of elegant experiments, Urban and Vogel demonstrated the regulation of *glmS* by GlmZ and GlmY. Through the use of primer extension assays, it was shown that the expression of either GlmZ or GlmY induces the production of monocistronic *glmS* mRNA, mimicking the effects seen in mutants in which *pcnB* or *yhbJ* have been deleted.[19,21] Following from this, a plasmid was developed which expressed a *glmS::gfp* fusion mRNA and was activated in the presence of GlmY. A colony of *E. coli* was developed which expressed this plasmid in combination with a control plasmid, a plasmid expressing GlmY or a plasmid expressing GlmZ. GlmZ was found to activate the fusion plasmid just as efficiently as GlmY, indicating to researchers that the fusion plasmid contained the region of *glmS* involved in regulation by GlmZ and GlmY. Through alignment of the GlmY and GlmZ RNAs, it was shown that the two sRNAs possess similar structure and sequence homology. It might be assumed that, because of this, the two are redundant. However, transformation of wild type *E. coli* or ΔGlmZ or ΔGlmY mutant strains with GlmY or GlmZ plasmids showed that GlmY cannot upregulate *glmS* independently of GlmZ. GlmY upregulated *glmS* in the wild-type and ΔGlmY mutants but not in ΔGlmZ mutants, however the same level of GlmY RNA was produced, indicating that the absence of GlmZ did not prevent the accumulation of GlmY RNA. GlmY and GlmZ were expressed in other mutants combining deletions of *glmY*, *glmZ*, *hfq* and *yhbJ*, and the results collectively showed that there is no GlmZ-independent path for GlmY to regulate *glmS*.

Following from this, it was predicted that GlmZ operated by unblocking a hairpin in the *glmS* mRNA that would otherwise block the Shine-Delgarno sequence. To investigate this, a *glmS::gfp* fusion plasmid was developed with a mutation in the GlmZ binding site that still allowed the formation of the hairpin. When GlmZ was introduced *glmS* was no longer activated. However, upon introduction of a mutated GlmZ plasmid, activation was restored, confirming the mechanism of activation. *In vitro* results also confirmed the activation of *glmS* by GlmZ.[22] In addition, it was shown that only full-length GlmZ and not the processed RNA form can activate *glmS*. The establishment of the role of GlmZ naturally introduced questions about

the role of GlmY in regulation of *glmS*. It was thought that GlmY increases the amount of full-length GlmZ in the cell. An arabinose-inducible P_{BAD} promoter was inserted into *E. coli* Δ*glmY* mutants and it was shown that GlmZ processing in the resulting recombinant cells was fully suppressed, and the levels of full-length GlmZ RNA and *glmS* in the cell consequently increased.[22] This suggested to researchers that GlmY is an essential component in the stabilisation of GlmZ. This was supported by the loss of *glmS::gfp* fusion activity in a *pcnBΔ1* Δg*lmY* double mutant, which revealed that *glmY* is required for *glmS* activation in *pcnBΔ1*. Also of interest was the fact that the half-life of GlmZ was reduced in the double mutant but restored upon introduction of ectopic GlmY. This confirmed that GlmY activates *glmS* by preventing GlmZ decay.[22] Finally, an RNA circularisation assay was used to investigate the level of polyadenylation of GlmZ, GlmY and *glmS* transcripts. Polyadenylated tails were found at the 3' end of the GlmY RNA, and not the GlmZ or *glmS* transcripts. In addition, none of the three RNAs were polyadenylated in *pcnBΔ1* mutant cells. Thus, researchers could confirm that the polyadenylation of the GlmY transcript by PAP I destabilises it and consequently leads to stabilisation of GlmZ and the expression of *glmS*.[22] An overview of this method of regulation is shown in Figure 3. This system is a prime example of how many factors can act hierarchically to regulate mRNA stability, and thus alter gene expression to meet the cell's needs.

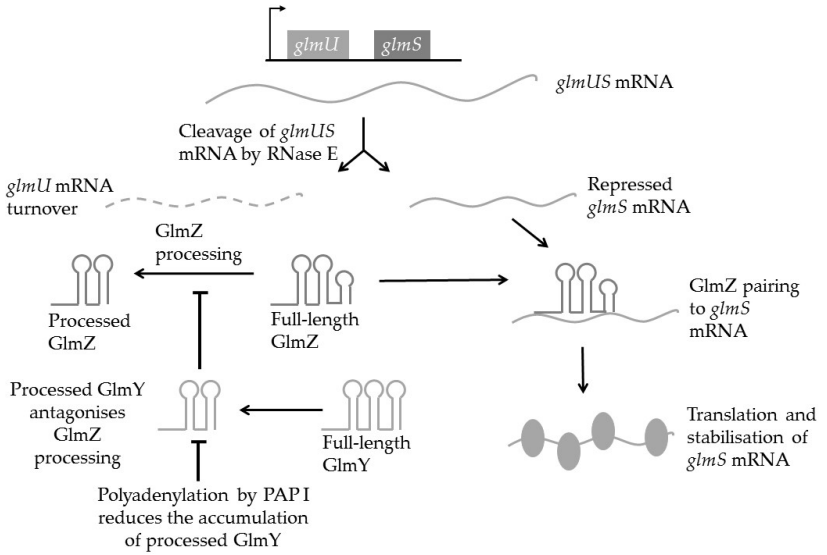


Figure 3: Regulation of *glmS* stability in *E. coli*. The transcription of *glmS* from the *glmUS* gene is shown, as is the processing of GlmZ by GlmY.and its downstream effect on *GlmS* stability
Figure adapted from Urban and Vogel, 2008

Conclusion

Bacteria are commonly thought to be the most successful life form on earth, due to their vast numbers, species versatility, and their ability to survive and thrive in a myriad of harsh environments, such as extreme temperature, pressure and pH. For this reason, prokaryotic gene expression has always been of great interest to researchers. Bacteria are often thought of as simple beings, consisting of a cell membrane and a loop of DNA, and thus are naturally thought to possess an equally simple gene expression system. However, as the example in this review has shown, prokaryotic gene expression is far from simple. Both on the transcription and translation level, an extensive network of components cooperate to ensure rapid adaptation to the environment. The recent research completed on translational regulation, in particular, revealed a multitude of tightly-regulated RNA-protein networks whose function it is to control mRNA stability and thus, translation. Indeed, it is mechanistic – a collection of components cooperating in such a structured, sometimes hierarchical, manner – however it is not the primitive gene expression system we would expect from such a seemingly simple organism. That being said, the very existence of some aspects of translational regulation, such as riboswitches, as discussed previously, provides some evidence in support of the RNA world hypothesis. Thus, the complex mRNA regulation systems employed by modern bacteria provide a very useful insight into the simplicity of the beginnings of life.

Regulation of mRNA stability as a method of controlling gene expression in bacteria has long been underestimated in the face of transcriptional regulation. Despite the extensive research carried out in recent years, it is evident that there are a lot more aspects to it which remain to be uncovered. The regulation of mRNA stability offers the possibility of a genetic regulatory network much more extensive than was previously thought, which would thus provide a clear insight into the cellular adaptation process.

References

1. Kushner SR. mRNA decay in Escherichia coli comes of age. *J Bacteriol*. 2002 Sep;184(17):4658–65; discussion 4657.

2. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell*. 2011 Sep 16;43(6):880–91.

3. Vogel J, Luisi BF. Hfq and its constellation of RNA. *Nat Rev Microbiol*. 2011;9(8):578–89.

4. Pedersen M, Nissen S, Mitarai N, Svenningsen S Lo, Sneppen K, Pedersen S. The Functional Half-Life of an mRNA Depends on the Ribosome Spacing in an Early Coding Region. *J Mol Biol*. 2011;407(1):35–44.

5. Carpousis AJ. The RNA Degradosome of Escherichia coli : An mRNA-Degrading Machine Assembled on RNase E. *Annu Rev Microbiol*. 2007 Oct;61(1):71–87.

6. Mackie GA. Ribonuclease E is a 5|[prime]|-end-dependent endonuclease. *Nature*. 1998 Oct 15;395(6703):720–4.

7. Celesnik H, Deana A, Belasco JG. Initiation of RNA Decay in Escherichia coli

by 5??? Pyrophosphate Removal. *Mol Cell.* 2007;27(1):79–90.
8. Py B, Higgins CF, Krisch HM, Carpousis AJ. A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. *Nature.* 1996 May 9;381(6578):169–72.

9. Miczak A, Kaberdin VR, Wei C, Lin-chao S, by Stanley Cohen CN. Proteins associated with RNase E in a multicomponent ribonucleolytic complex (DnaK/enolase/RNA helicase). *Biochemistry.* 1996;93:3865–9.

10. Feng Y, Cohen SN. Unpaired terminal nucleotides and 5′ monophosphorylation govern 3′ polyadenylation by Escherichia coli poly(A) polymerase I. *Proc Natl Acad Sci.* 2000 Jun 6;97(12):6415–20.

11. Yehudai-Resheff S, Schuster G. Characterization of the E.coli poly(A) polymerase: nucleotide specificity, RNA-binding affinities and RNA structure dependence. *Nucleic Acids Res*. 2000 Mar 1;28(5):1139–44.

12. Mohanty BK, Kushner SR. Analysis of the function of Escherichia coli poly(A) polymerase I in RNA metabolism. *Mol Microbiol.* 1999 Dec;34(5):1094–108.

13. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*. 2004 Mar 18;428(6980):281–6.

14. Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, *et al*. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci U S A*. 2004 Apr 27;101(17):6421–6.

15. Even S, Pellegrini O, Zig L, Labas V, Vinh J, Bréchemmier-Baey D, *et al*. Ribonucleases J1 and J2: two novel endoribonucleases in B.subtilis with functional homology to E.coli RNase E. *Nucleic Acids Res.* 2005;33(7):2141–52.

16. Winkler WC, Breaker RR. Regulation of Bacterial Gene Expression by Riboswitches. *Annu Rev Microbiol*. 2005 Oct;59(1):487–517.

17. Collins JA, Irnov I, Baker S, Winkler WC. Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev.* 2007 Dec 15;21(24):3356–68.
18. Plumbridge J. Co-ordinated regulation of amino sugar biosynthesis and degradation: the NagC repressor acts as both an activator and a repressor for the transcription of the glmUS operon and requires two separated NagC binding sites. *EMBO J.* 1995 Aug 15;14(16):3958–65.

19. Kalamorz F, Reichenbach B, März W, Rak B, Görke B. Feedback control of glucosamine-6-phosphate synthase *GlmS* expression depends on the small RNA GlmZ and involves the novel protein YhbJ in Escherichia coli. *Mol Microbiol*. 2007 Sep;65(6):1518–33.

20. Urban JH, Papenfort K, Thomsen J, Schmitz RA, Vogel J. A Conserved Small RNA Promotes Discoordinate Expression of the glmUS Operon mRNA to Activate *GlmS* Synthesis. Vol. 373, *Journal of Molecular Biology.* 2007.

21. Joanny G, Le Derout J, Bréchemier-Baey D, Labas V, Vinh J, Régnier P, *et al*. Polyadenylation of a functional mRNA controls gene expression in *Escherichia coli. Nucleic Acids Res*. 2007;35(8):2494–502.

22. Urban JH, Vogel J. Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol*. 2008;6(3):0631–42.

TS
SR

Chemistry

TS
SR

Letter from the Editor

People in science often speak of "impact", by which they mean the economic, societal and other benefits provided by scientific research. This volume of TSSR-Chemistry is one that demonstrates the impact that chemistry has on our world. It cements the long held view that chemistry is the central science; the most extensively connected of all. The research discussed by this year's authors is directly connected to the biggest issues facing humanity, from cancer, to climate change, energy production, energy storage, and much more. The reviews presented here highlight the interest and determination of Trinity students to tackle these big problems. We hope that their reviews will inspire others to follow suit.

Of all the molecules, in all the world, in all the universe, there is no one more famous than CO_2. We are releasing it into the atmosphere on an unprecedented scale. Eoin Dolan's review describes the on-going efforts to reduce CO_2 to methanol: a stable liquid that could be used as both a fuel and feedstock chemical. Cancer, a group of diseases, that has impacted the lives of so many, is being intensively investigated by scientists. Mark Berney's review demonstrates the role chemists have to play in those investigations. The review discusses current research into the quantification of DNA modifications, which have been directly implicated in cancer. The review highlights the growing role chemistry has to play in solving biological problems. The field of nanoscience seems destined to have a major impact on our world. Dáire Brady's review discusses how nanoscience is pushing lithium-ion batteries to the limit. This paper provides a glimpse of the on-going efforts to tackle the energy problems of the future. Energy is again the topic of discussion in Tim Ryan's review on redox flow batteries. This is a promising field of research, which may well be a solution to our energy storage problems.

As is always the case with large projects like the TSSR, its success depended on the help of many people. I would thus like to take this opportunity to thank those who were instrumental in another successful volume of the TSSR. Firstly I would like to acknowledge the academic advisor - Dr. Mike Southern - who generously offered up his time to provide feedback and support for the journal, and for that I am extremely grateful. On behalf of the TSSR team and the student authors, I would like to thank the academic reviewers – Dr. Maria Daniela Angione, Dr. Michael Lyons, Dr. Niall McEvoy, Dr. Joanna McGouran, Maria O'Brien, and Dr. Wolfgang Schmitt – for their time and effort in reviewing the submissions. Finally, I would like to extend my thanks and appreciation to all who submitted papers to TSSR-Chemistry this year. It was a pleasure to read about the various topics that interest fellow students at Trinity. I appreciate the hard work and time taken outside of coursework to contribute to the TSSR. I am

confident that all those who submitted will be successful in their future scientific endeavours.

In summary, this year's volume of TSSR-Chemistry was made possible by the efforts of many people including the TSSR team, the student authors, and the academic advisors. The papers presented here describe topics that tackle the biggest problems facing humanity. They affirm the position of chemistry as the central science. Indeed, Linus Pauling's words are pertinent: "Every aspect of the world today – even politics and international relations – is affected by chemistry".

Stephen Byrne
Chemistry Editor
Trinity Student Scientific Review 2017

Epigenetic Cytosine Modifications and their Detection

Mark Berney
Senior Sophister
Chemistry

Methylation of cytosine at the 5-position is a common epigenetic marker in mammalian DNA, and plays an important role in regulating gene expression. Oxidised derivatives of 5 methylcytosine have recently been discovered. As well as being intermediates in an active demethylation pathway, some of these oxidised derivatives seem to function as epigenetic markers in their own right. There is a pressing need to learn more about these cytosine modifications, due to the role epigenetic markers play in diseases such as cancer. Research in this area has been hampered however by the difficulty associated with detecting cytosine modifications with a high degree of accuracy and sequence specificity. This review will briefly summarise the biochemistry of cytosine modifications, and discuss new and established methods for their detection.

Introduction

DNA codes for all the proteins necessary for life, but the DNA sequence is not the sole determinant of all the phenotypic traits of cells and organisms. Transcription of DNA is tightly regulated by epigenetic modifications, which play an important role in controlling when and where specific genes are expressed.[1] Epigenetic modifications are implicated in important processes such as cellular differentiation, and also play a role in various diseases such as cancer.[2] Epigenetic control of gene expression is often mediated through covalent modification of DNA itself, or of the histone proteins around which DNA is wrapped in the nucleus. These modifications do not result in changes to the DNA sequence, but can affect the binding of certain proteins to the DNA, including transcription factors and proteins which regulate chromatin structure.[3–5]

In mammals and many other eukaryotes, the most common epigenetic covalent modification of DNA itself is methylation of cytosine (C) at the 5-position. Such modifications can be inherited, but can also be enzymatically introduced and removed in response to stimuli.[6] 5-Methylcytosine (mC) is introduced by the methylation of cytosine residues by DNA-methyltransferases (DNAMTs) with an S adenosylmethionine cofactor. 5 Methylcytosine is most often found in the context of 5' cytosine-phosphate-guanine-3' (CpG) dinucleotides. CpG islands, or regions containing a high density of CpG sequences, are found in about 72% of promoters in the human genome.[7] Methylation of these regions causes transcriptional inactivation of the associated gene.[6] One of the roles which epigenetics plays in cancer is the silencing of tumour suppressor genes by the methylation of CpG islands in their promoters.[8] DNAMT inhibitors can thus be effective anti-cancer drugs.[9]

It was previously believed that demethylation of mC occurred only by passive dilution, the replication of cells without maintenance of methylation patterns. However the recent discovery of oxidised forms of mC has revealed a new active demethylation pathway, summarised in Figure 1. 5-Methylcytosine can be converted to 5 hydroxymethycytosine (hmC) by the ten-eleven translocation dioxygenase (TET) family of enzymes, most prominently by TET1. A dioxygen molecule is transferred to α-ketoglutarate and mC *via* reactive Fe(III)/Fe(IV) intermediates to give succinate and hmC.[10] 5 Hydroxymethylcytosine can be further oxidised by TET1 to 5 formylcytosine (fC) and 5 carboxylcytosine (caC),[11] both of which can be excised by thymine DNA glycosylase (TDG) to give an a basic site (a sugar unit with no nucleobase attached).[12] Unmethylated cytosine can then be restored by base excision repair (BER).

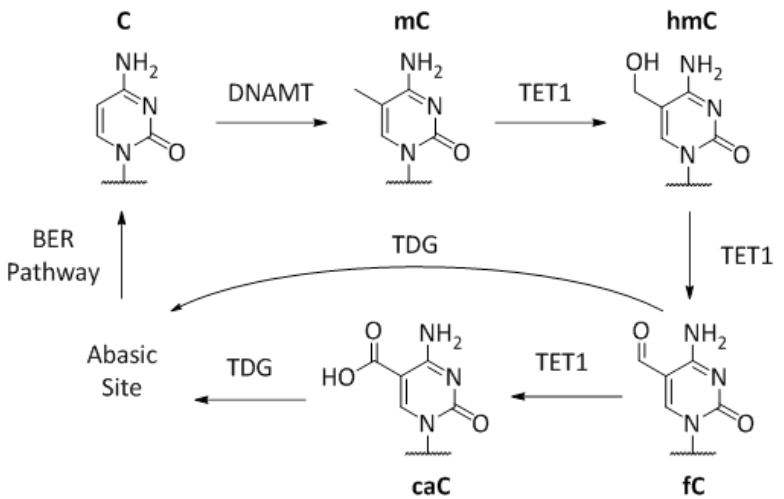


Figure 1: Active Demethylation of Cytosine

There is some evidence that oxidised mC derivatives serve not just as intermediates in this demethylation pathway, but also as epigenetic markers in their own right.[13] Modifications at the cytosine 5-position protrude into the major groove of DNA and are thus available for interaction with DNA-binding molecules. Differences in their hydrogen-bonding properties, as well as steric differences, allow differentiation between them. The formyl group in fC is held rigidly in the plane of the cytosine ring due to an intramolecular hydrogen bond with the amino group at the 4-position.[13] On the other hand the hydroxy group in hmC has been found in two different conformations in crystal structures[14] and thus appears to rotate freely. This exaggerates what would otherwise be a slight difference in their steric properties, allowing for easier discrimination between these modifications by DNA-binding proteins. Cytosine modifications have even been shown to distort the double-helix and alter the conformation of DNA when present at high density,[15] and can influence the binding of DNA intercalators.[16]

Some hmC modifications, particularly those at promoters and at poised and active enhancers, appear to be stable and not subject to further oxidation by TET1.[17] 5 Hydroxymethylcytosine is also enriched in exons and near transcriptional start sites.[18] 5 Hydroxmethylcytosine has been found to be especially abundant in the brain, particularly in purkinje neurons,[19] and in embryonic stem cells. Certain cancer cell types on the other hand have lower than normal hmC content.[19] All this suggests that hmC plays a role in gene regulation and that cell-type specific regulation of TET dioxygenases may be an important mediator of epigenetic control.[13]

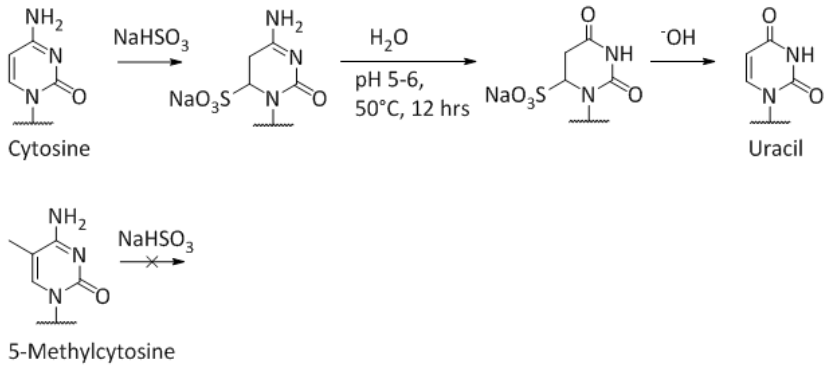
5-Formylcytosine appears to be a semi-permanent marker in some cases as well, as TDG removes only 50% of fC at specific genomic sites.[20] 5-Formylcytosine is known to preferentially occur at poised enhancers in mouse embryonic stem cells,[21] and several proteins have been identified, including transcription factors and chromatin regulators, which have a strong preference for binding to regions containing fC.[22] 5-Formylcytosine is also found in mitochondrial tRNA molecules, where it modulates codon anticodon interactions.[23]

Oxidised mC derivatives thus seem to play an important biological role, not only in demethylation pathways, but are not yet fully understood.

Chemical Methods for the Detection of Cytosine Modifications

Mass spectrometry based methods have been used to study genome wide levels of epigenetic cytosine modifications.[24–29] The sequence specific detection of cytosine modifications can be achieved using restriction enzymes,[30] but the most widely used strategy is the chemical derivatisation of DNA using sodium bisulfite in a

process known as bisulfite sequencing, illustrated in Scheme 1. Treatment of DNA with sodium bisulfite leads to deamination of C to uracil (U), but not of mC to thymine (T).



Scheme 1: Bisulfite Sequencing

Treatment with bisulfite therefore translates epigenetic information into a change in the sequence of canonical nucleobases, which can be detected using PCR followed by DNA sequencing, in which C will be read as U and mC will be read as C. Simple bisulfite sequencing cannot however distinguish the oxidised derivatives of mC. Cytosine 5 methylsulfonate (CMS) is formed on treatment of hmC with bisulfite, and since CMS shows the same base-pairing selectivity for guanine, it reads as C.[31] Bisulfite converts fC and caC to U, after deformylation or decarboxylation, and therefore neither of these are distinguishable from unmodified C by this method.[13]

This limitation can be overcome by the use of additional chemical derivatisation steps before bisulfite treatment. Oxidation of hmC to fC with potassium perruthenate, or reduction of fC to hmC with sodium borohydride, followed by bisulfite sequencing, allows fC and hmC levels to be determined by comparing oxidised/reduced samples with untreated ones.[32] Another additional derivatisation step that has been employed to detect hmC, is the use of β-glucosyltransferase to glucosylate hmC. Subsequent treatment of the DNA with TET1 oxidises all 5-modified cytosine residues to caC, except hmC which is now protected. In bisulfite sequencing hmC is then the only base that reads as C.[13]

Although bisulfite sequencing is currently considered the gold standard for detecting epigenetic cytosine modifications, it has a number of drawbacks. About 95% of the DNA is destroyed on treatment with bisulfite, and so a large sample of DNA is needed.[33] Incomplete conversion of C to U can lead to errors,[34] and the use of several chemical manipulations of the sample increases the chances for bias or contamination.[35] Bisulfite sequencing is also labour intensive, since several steps are required.

A number of other chemical derivatisation strategies have been developed to detect mC, but have not yet been extended to the analysis of oxidised mC derivatives. Treatment of DNA containing mC with osmium tetroxide causes oxidation of mC to give an osmate complex, while unmodified C does not react.[36] A bipyridine ligand modified with a linker attached to a fluorescent or electrochemically active group can coordinate to the osmate and allow detection of mC.[37] Detection of mC by selective oxidation has also been achieved using V_2O_5, or $NaIO_4$ with LiBr.[38] Direct electrochemical oxidation has also been used.[39,40] Derivatisation of DNA with reagents such as N-halogeno-N-sodiobenzenesulfonamide[41] or O allylhydroxylamine[42] have also been used to detect mC.

β-Glucosyltransferase has been used outside the context of bisulfite sequencing for detection of hmC. After glucosylation of hmC, the DNA may be treated with sodium periodate to oxidise the glucose, and the resulting aldehydes then allow for attachment of biotin probes *via* the formation of oxime linkages. Biotin tags allow for detection of DNA fragments containing hmC using streptavidin in a pull-down assay.[18] Alternatively a glucose moiety containing an azide group can be enzymatically attached to hmC, and a biotin group then added using a copper catalysed 1,3-dipolar cycloaddtion.[43]

A promising recently published example of a chemical derivatisation strategy for the detection of fC is the labelling of DNA with trimethylindole derivatives, which react with fC to produce hemicyanine-like chromophores.[44] Quantitative measurement of fC levels can then be achieved by measurement of the intensity of the fluorescence emission of the sample. Site specific detection of fC can also be achieved using a primer extension assay, since the hemicyanine-modified nucleobases act as a roadblock to Klenow DNA polymerase, which can usually bypass fC.

Approaches Based on Non-Covalent Interactions with Molecules that Recognise DNA

The main alternative to these chemical derivatisation strategies is to use DNA-binding proteins that can recognise epigenetic cytosine bases. An early method for detecting mC was based on the use of restriction endonucleases that include CpG dinucleotides in their recognition sequences, and do not cleave DNA that is methylated at cleavage/recognition sites.[45] A more common strategy for the detection of mC is the use of antibodies in a methylated DNA immunoprecipitation assay (MeDIP). DNA is first fragmented, typically by sonication, and then denatured. The resulting single strand fragments which contain mC are bound by monoclonal mC antibodies. The fragments bound to antibodies can then be separated using magnetic beads coated with mouse antibodies that bind the monoclonal mC antibodies. This allows the fragments to be analysed. Antibodies specific for each modified form of cytosine have been used in this kind of assay,[13]

as have antibodies specific for CMS.[18] However this method has several drawbacks compared to bisulfite sequencing methods. It is not quantitative, and the resolution is dependent on the size of the DNA fragments, as any number of mC residues in the fragment will lead to a positive signal. Methyl-CpG-binding domains of MeCP2 proteins can also be used to bind mC and hmC without programmable sequence specificity in the context of CpG dinucleotides.[13] J-binding protein 1 can recognise glucosylated hmC.[13]

Transcription activator-like effectors (TALEs), a type of protein found in Xanthomonas bacteria, have been used to recognise cytosine derivatives with programmable sequence specificity.[46] The sequence specificity of TALEs is derived from the DNA-binding domain, which consists of repeat units 34 amino acids in length. The 12th and 13th amino acids in each repeat unit are found in a loop between two α-helices and constitute the repeat variable diresidue (RVD). The four naturally occurring RVDs recognise the four canonical nucleobases through hydrogen-bonding interactions. TALEs have been engineered which contain additional mutant RVDs that recognise modified cytosine nucleobases. So far, only mC and hmC have been selectively detected in this way.

Another method which shows promise for detecting cytosine modifications is nanopore sequencing. Proteins forming nanopores in a barrier separating two compartments filled with electrolyte allow ions to flow through them when an electric potential is applied. DNA can also migrate through the nanopores, and in doing so modulates the ionic current in a way that depends on the structure of the nucleotides present in the nanopore. Monitoring of the ionic current over time therefore allows sequencing of the DNA strand as it moves through the pore. Controlling the kinetics of the DNA translocation through the pore improves the accuracy of the sequencing. This has been achieved by the use of the highly processive phi29 DNA polymerase, which acts as a cap on the pore and slowly threads DNA through. A nanopore sequencing method using a mutant form of the MspA porin protein found in Mycobacterium smegmatis has been shown to allow discrimination between C, mC, hmC, fC, and caC.[47] The accuracy of nanopore sequencing methods is comparatively low however (92-98% accurate in reference oligonucleotides of known sequence), but engineering of the pore forming proteins may lead to further improvements.

The Use of Oligonucleotides for Detecting Cytosine Modifications

Oligonucleotide probes offer an advantage over protein based probes such as TALEs since they are more cost-effective, as they can be more readily prepared using phosphoramidite building blocks in an automated solid-phase synthesis process. A large number of modified phosphoramidites have been developed to extend this method to the synthesis of oligonucleotides containing non-native

functional groups.[48–51] Oligonucleotide probes also offer easily tuneable selectivity using probe sequences which are complementary to target sites.

A DNA templated photoligation has been used to detect 5-methylcytosine.[52] A probe strand containing a terminal 5-vinyl-2'-deoxyuridine with a hydrophobic group, undergoes a [2+2] cycloaddition with the carbon-carbon double bond of mC upon irradiation at 366 nm. Measurement of the fluorescence emission of the product allows easy detection. Reaction with mC is much more efficient than reaction with unmodified C due to a favourable hydrophobic interaction, illustrated by the green arrow in Figure 2, between the methyl group of mC and the various hydrophobic moieties which were tested in the 5 vinyl 2' deoxyuridine residue.

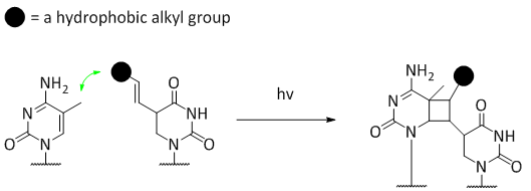


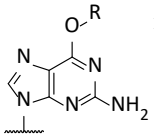
Figure 2: A [2+2] Cycloaddition Between mC and a 5-Vinyl-2'-Deoxyuridine Residue Promoted by a Hydrophobic Interaction

As well as the example discussed above, another DNA templated light-activated reaction has been used to detect mC. An oligonucleotide probe tethered to a sensitizing 2 methyl 1,4 naphthoquinone chromophore causes one-electron oxidation of mC in a complementary strand upon irradiation. Treatment with piperidine then leads to selective oxidative strand cleavage at mC. Oligonucleotides in which mC is replaced with one of the four canonical nucleobases are much less susceptible to strand cleavage.[53]

Other oligonucleotide probes based on fluorescent detection have also been reported. A DNA probe containing an anthracene fluorophore has been shown to discriminate between all four canonical nucleobases as well as mC through differences in the response of fluorescence emission properties to formation of a duplex with the strand under analysis.[54] An oligonucleotide containing a fluorescein moiety at the 5'-end and a dabsyl quencher group at the 3'-end has been used to distinguish between C and mC. The oligonucleotide probe first anneals to the region under analysis. A restriction enzyme is then used to digest the resulting complex, leading to strong fluorescence when C is present as the part of the probe containing the fluorescein is separated from the part containing the quencher. When mC is present however the restriction enzyme cannot digest the sample and only weak fluorescence is observed.[55] Beyond the use of fluorescence detection, an electrochemically active ferrocene acetic acid group has been conjugated to the 3'-end of a probe oligonucleotide to detect mC.[56]

Modified nucleotides have been used to discriminate between C and mC in primer extension assays, as O^6-modified 2' deoxyguanosine derivatives are incorporated opposite C or mC with different efficiencies.[57]

Modified nucleotides have also been incorporated into triplex forming oligonucleotides to detect modified cytosines. It has been observed that triplex forming oligonucleotides containing a synthetic N-methylpyrrolocytosine base show significantly lower triplex melting temperatures when bound to strands which have mC rather than C in a CpA sequence.[58]



R = methyl, ethyl, propyl or isopropyl

Figure 3: O^6-Modified 2'-Deoxyguanosine Derivatives for Detecting mC.

Carell *et al*. have recently reported the first use of an oligonucleotide probe for the general and reliable detection of an oxidised derivative of mC.[20] This was done using an oligonucleotide probe modified to contain a hydroxylamine moiety to allow for the formation of an oxime crosslinkage with the formyl group of an fC residue in a complementary strand.

Conclusion

No new general method for the sequence specific detection of mC, hmC, fC, and caC has yet been developed to overcome the drawbacks associated with bisulfite sequencing. Nanopore sequencing shows promise however, if the accuracy can be improved. The use of oligonucleotide probes for detecting cytosine modifications is an emerging area which offers potential for the development of probes to detect epigenetic modifications at specific sites without the need for DNA sequencing. Oligonucleotide probes could potentially be immobilised on a microarray for the convenient screening of DNA samples, or used in vivo in a manner similar to established fluorescent *in situ* hybridisation (FISH) techniques.[59]

Further advancements in the detection of cytosine modifications will facilitate research on their role in cancer and certain developmental diseases, potentially leading to new treatments. Analysis of DNA is also important in diagnostics and forensics, and sequence specific probes for epigenetic cytosine derivatives may find applications in these areas as well.

Acknowledgements

The author would like to thank Prof. Joanna McGouran for her help and advice. This work was adapted from a thesis entitled "The Butterfly Effect: Site Selective Detection of Epigenetic Modifications" submitted for the course CH4119.

References

1. Jaenisch, R., and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Gen* 33, 245–254.

2. Kinnaird, A., Zhao, S., Wellen, K. E., and Michelakis, E. D. (2016) Metabolic control of epigenetics in cancer. *Nat Rev Cancer* 16, 694–707.

3.Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R. W., Deaton, A., Andrews, R., James, K. D., Turner, D. J., Illingworth, R., and Bird, A. (2010) CpG islands influence chromatin structure *via* the CpG-binding protein Cfp1. *Nature* 464, 1082–1086.

4. Allis, C. D., and Jenuwein, T. (2016) The molecular hallmarks of epigenetic control. *Nat Rev Genet* 17, 487–500.

5. Cook, P. C., Owen, H., Deaton, A. M., Borger, J. G., Brown, S. L., Clouaire, T., Jones, G.-R., Jones, L. H., Lundie, R. J., Marley, A. K., Morrison, V. L., Phythian-Adams, A. T., Wachter, E., Webb, L. M., Sutherland, T. E., Thomas, G. D., Grainger, J. R., Selfridge, J., McKenzie, A. N. J., Allen, J. E., Fagerholm, S. C., Maizels, R. M., Ivens, A. C., Bird, A., and MacDonald, A. S. (2015) A dominant role for the methyl-CpG-binding protein Mbd2 in controlling Th2 induction by dendritic cells. *Nat. Commun.* 6, 6920.

6. Law, J. a, and Jacobsen, S. E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–20.

7. Saxonov, S., Berg, P., and Brutlag, D. L. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. U. S. A.* 103, 1412–1417.

8. Esteller, M. (2002) CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* 21, 5427–40.

9. Jordheim, L. P., Durantel, D., Zoulim, F., and Dumontet, C. (2013) Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat. Rev. Drug Discov.* 12, 447–464.

10. Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y. G., Zhu, J., Wang, P., and Xu, Y. (2013) Crystal Structure of TET2-DNA Complex: Insight into TET-Mediated 5mC Oxidation. *Cell* 155, 1545–1555.

11. Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. a, He, C., and Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* (80-.). 333, 1300–3.

12. He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C.-X., Zhang, K., He, C., and Xu, G.-L. (2011) Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Sci. (Washington, DC, United States)* 333, 1303–1307.

13. Kubik, G., and Summerer, D. (2015) Deciphering Epigenetic Cytosine Modifications by Direct Molecular Recognition. *ACS Chem. Biol.* 10, 1580–1589.

14. Szulik, M. W., Pallan, P. S., Nocek, B., Voehler, M., Banerjee, S., Brooks, S.,

Joachimiak, A., Egli, M., Eichman, B. F., and Stone, M. P. (2015) Differential stabilities and sequence-dependent base pair opening dynamics of Watson-crick base pairs with 5-hydroxymethylcytosine, 5-formylcytosine, or 5-carboxylcytosine. *Biochemistry* 54, 1294–1305.

15. Raiber, E.-A., Murat, P., Chirgadze, D. Y., Beraldi, D., Luisi, B. F., and Balasubramanian, S. (2015) 5-Formylcytosine alters the structure of the DNA double helix. *Nat. Struct. Mol. Biol.* 22, 44–49.

16. Zhang, N., Lin, C., Huang, X., Kolbanovskiy, A., Hingerty, B. E., Amin, S., Broyde, S., Geacintov, N. E., and Patel, D. J. (2005) Methylation of Cytosine at C5 in a CpG Sequence Context Causes a Conformational Switch of a Benzo[a] pyrene diol epoxide-N2-guanine Adduct in DNA from a Minor Groove Alignment to Intercalation with Base Displacement. *J. Mol. Biol.* 346, 951–965.

17. Pastor, W. A., Aravind, L., and Rao, A. (2013) TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.* 14, 341–356.

18. Pastor, W., Pape, U., Huang, Y., and Henderson, H. (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473, 394–397.

19. Kriaucionis, S., and Heintz, N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324, 929–930.

20. Su, M., Kirchner, A., Stazzoni, S., Müller, M., Wagner, M., Schröder, A., and Carell, T. (2016) 5-Formylcytosine Could Be a Semipermanent Base in Specific Genome Sites. Angew. *Chemie Int. Ed.* 55, 11797–11800.

21. Song, C. X., Szulwach, K. E., Dai, Q., Fu, Y., Mao, S. Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H., Gao, J., Liu, P., Li, L., Xu, G. L., Jin, P., and He, C. (2013) Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* 153, 678–691.

22. Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.-A., Bachman, M., Booth, M. J., Andrews, S., Balasubramanian, S., and Reik, W. (2013) A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. *Genome Biol.* 14, R119.

23. Lusic, H., Gustilo, E. M., Vendeix, F. A. P., Kaiser, R., Delaney, M. O., Graham, W. D., Moye, V. A., Cantara, W. A., Agris, P. F., and Deiters, A. (2008) Synthesis and investigation of the 5-formylcytidine modified, anticodon stem and loop of the human mitochondrial tRNAMet. *Nucleic Acids Res.* 36, 6548–6557.

24. Pfaffeneder, T., Hackner, B., Truß, M., Münzel, M., Müller, M., Deiml, C. A., Hagemeier, C., and Carell, T. (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA. Angew. *Chemie - Int. Ed.* 50, 7008–7012.

25. Schiesser, S., Hackner, B., Pfaffeneder, T., Müller, M., Hagemeier, C., Truss, M., and Carell, T. (2012) Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. Angew. *Chemie - Int. Ed.* 51, 6516–6520.

26. Friso, S., Choi, S.-W., Dolnikowski, G. G., and Selhub, J. (2002) A Method to Assess Genomic DNA Methylation Using High-Performance Liquid Chromatography/ Electrospray Ionization Mass Spectrometry. *Anal. Chem.* 74, 4526–4531.

27. Münzel, M., Globisch, D., Brückl, T., Wagner, M., Welzmiller, V., Michalakis, S., Müller, M., Biel, M., and Carell, T. (2010) Quantification of the sixth DNA base hydroxymethylcytosine in the brain. Angew. *Chemie - Int. Ed.* 49, 5375–5377.

28. Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M., and Carell, T. (2010) Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates. *PLoS One* (Croft, A. K., Ed.)

5, e15367.

29. Wagner, M., Steinbacher, J., Kraus, T. F. J., Michalakis, S., Hackner, B., Pfaffeneder, T., Perera, A., Müller, M., Giese, A., Kretzschmar, H. a., and Carell, T. (2015) Age-Dependent Levels of 5-Methyl-, 5-Hydroxymethyl-, and 5-Formylcytosine in Human and Mouse Brain Tissues. Angew. *Chemie - Int.* Ed. 54, 12511–12514.

30. Booth, M. J., Raiber, E., and Balasubramanian, S. (2015) Chemical Methods for Decoding Cytosine Modi fi cations in DNA. *Chem. Rev.* 2240–2254.

31. Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R., and Rao, A. (2010) The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing. *PLoS One 5*, e8888.

32. Booth, M. J., Marsico, G., Bachman, M., Beraldi, D., and Balasubramanian, S. (2014) Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat. Chem.* 6, 435–40. 33Grunau, C., Clark, S. J., and Rosenthal, A. (2001) Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* 29, e65–e65.

34. Genereux, D. P., Johnson, W. C., Burden, A. F., Stöger, R., and Laird, C. D. (2008) Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. *Nucleic Acids Res.* 36, e150–e150.

35. Clark, S. J., Statham, A., Stirzaker, C., Molloy, P. L., and Frommer, M. (2006) DNA methylation: Bisulphite modification and analysis. *Nat. Protoc.* 1, 2353–2364.

36. Okamoto, A., Tainaka, K., and Kamei, T. (2006) Sequence-selective osmium oxidation of DNA: efficient distinction between 5-methylcytosine and cytosine. *Org. Biomol. Chem*. 4, 1638.

37. Tanaka, K., Tainaka, K., Kamei, T., and Okamoto, A. (2007) Direct labeling of 5-methylcytosine and its applications. *J. Am. Chem. Soc.* 129, 5612–5620.

38. Bareyt, S., and Carell, T. (2008) Selective detection of 5-methylcytosine sites in DNA. Angew. *Chemie - Int. Ed*. 47, 181–184.

39. Kato, D., Sekioka, N., Ueda, A., Kurita, R., Hirono, S., Suzuki, K., and Niwa, O. (2008) A Nanocarbon Film Electrode as a Platform for Exploring DNA Methylation. *J. Am. Chem. Soc.* 130, 3716–3717.

40. Wang, P., Mai, Z., Dai, Z., and Zou, X. (2010) Investigation of DNA methylation by direct electrocatalytic oxidation. *Chem. Commun. (Camb)*. 46, 7781–3.

41. Wang, T., Hong, T., Tang, T., Zhai, Q., Xing, X., Mao, W., Zheng, X., Xu, L., Wu, J., Weng, X., Wang, S., Tian, T., Yuan, B., Huang, B., Zhuang, L., and Zhou, X. (2013) Application of N-Halogeno-N-sodiobenzenesulfonamide Reagents to the Selective Detection of 5-Methylcytosine in DNA Sequences. *J. Am. Chem. Soc.* 135, 1240–1243.

42. Münzel, M., Lercher, L., Müller, M., and Carell, T. (2010) Chemical discrimination between dC and 5MedC *via* their hydroxylamine adducts. *Nucleic Acids Res.* 38.

43. Song, C.-X., Szulwach, K. E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.-H., Zhang, W., Jian, X., Wang, J., Zhang, L., Looney, T. J., Zhang, B., Godley, L. A., Hicks, L. M., Lahn, B. T., Jin, P., and He, C. (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotech* 29, 68–72.

44. Samanta, B., Seikowski, J., and Höbartner, C. (2016) Fluorogenic Labeling of 5-Formylpyrimidine Nucleotides in DNA and RNA. Angew. *Chemie - Int. Ed*. 55, 1912–1916.

45. Kaput, J., and Sneider, T. W. (1979) Methylation of somatic vs germ cell DNAs analyzed by restriction endonuclease digestions. *Nucleic Acids Res*. 7, 2303–2322.

46. Kubik, G., and Summerer, D. (2016) TALEored Epigenetics: A DNA-Binding Scaffold for Programmable Epigenome

Editing and Analysis. *ChemBioChem* 17, 975–980.

47. Wescoe, Z. L., Schreiber, J., and Akeson, M. (2014) Nanopores discriminate among five C5-cytosine variants in DNA. *J. Am. Chem. Soc.* 136, 16582–16587.

48. Beaucage, S. L., and Iyer, R. P. (1993) The Functionalization of Oligonucleotides *Via* Phosphoramidite Derivatives. *Tetrahedron* 49, 1925–1963.

49. Jl, Y., Bannwarth, W., and Luu, B. (1990) Application of the phosphoramidite-phosphite triester approach for the synthesis of combinations between oxygenated sterols and nucleoside analogues linked by phosphodiester bonds. *Tetrahedron* 46, 487–502.

50. Pon, R. T. (1991) A long chain biotin phosphoramidite reagent for the automated synthesis of 5′-biotinylated oligonucleotides. *Tetrahedron Lett.* 32, 1715–1718.

51. Stetsenko, D. A., and Gait, M. J. (2000) New phosphoramidite reagents for the synthesis of oligonucleotides containing a cysteine residue useful in peptide conjugation. Nucleosides. *Nucleotides Nucleic Acids* 19, 1751–1764.

52. Ogino, M., Taya, Y., and Fujimoto, K. (2009) Detection of methylcytosine by DNA photoligation *via* hydrophobic interaction of the alkyl group. *Org. Biomol. Chem*. 7, 3163.

53. Tanabe, K., Yamada, H., and Nishimoto, S. I. (2007) One-electron photooxidation and site-selective strand cleavage at 5-methylcytosine in DNA by sensitization with 2-methyl-1,4-naphthoquinone-tethered oligonucleotides. *J. Am. Chem. Soc.* 129, 8034–8040.

54. Duprey, J.-L. H. A., Zhao, Z., Bassani, D. M., Manchester, J., Vyle, J. S., and Tucker, J. H. R. (2011) Detection of DNA base variation and cytosine methylation at a single nucleotide site using a highly sensitive fluorescent probe. *Chem. Commun.* 47, 6629.

55. Okamoto, A., Tanabe, K., and Saito, I. (2002) Site-specific discrimination of cytosine and 5-methylcytosine in duplex DNA by peptide nucleic acids. *J. Am. Chem. Soc.* 124, 10262–10263.

56. Liu, S., Wu, P., Li, W., Zhang, H., and Cai, C. (2011) An electrochemical approach for detection of DNA methylation and assay of the methyltransferase activity. *Chem. Commun. (Camb)*. 47, 2844–6.

57. Von Watzdorf, J., Leitner, K., and Marx, A. (2016) Modified Nucleotides for Discrimination between Cytosine and the Epigenetic Marker 5-Methylcytosine. Angew. *Chemie - Int. Ed.* 55, 3229–3232.

58. Johannsen, M. W., Gerrard, S. R., Melvin, T., and Brown, T. (2014) Triplex-mediated analysis of cytosine methylation at CpA sites in DNA. *Chem. Commun*. (Camb). 50, 551–3.

59. Hu, L., Ru, K., Zhang, L., Huang, Y., Zhu, X., Liu, H., Zetterberg, A., Cheng, T., and Miao, W. (2014) Fluorescence in situ hybridization (FISH): an increasingly demanded tool for biomarker research and personalized medicine. *Biomark. Res.* 2, 3.

Recently Developed Catalysts for CO_2 Reduction to Methanol

Eoin Dolan
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

The environmental impact of elevated levels of CO_2 in the atmosphere has led to interest in capturing the gas and converting it into useful products. Hydrogenation of CO_2 to give methanol gives one potentially attractive use for this greenhouse gas. Methanol produced in this way provides both a carbon neutral biofuel and an efficient means of hydrogen storage and transportation. However current catalysts for producing methanol in this way, based on Cu and ZnO with metal oxide supports, require high concentrations of CO obtained from fossil fuels.[1] They also experience rapid deactivation by sintering in the temperature range of the reaction, show poor selectivity for methanol and are readily poisoned by water which builds up over the course of the reaction. Modifications to conventional catalysts give good improvements in these areas but a number of new types of catalyst have already demonstrated superior performance despite having undergone relatively little optimisation. This review will discuss some of the more promising catalysts based on alternative metals and support structures that might one day help to make the large scale recycling of CO_2 from the atmosphere a reality.

Introduction

With atmospheric CO_2 set to reach twice what it was before the industrial revolution by 2100,[2] increasing efforts are being made to move towards renewable alternatives to fossil fuels. Hydrogen is an attractive option due to the potentially limitless supply from electrolysis of seawater however it does present significant difficulties due to its volatile and gaseous nature, requiring high pressure facilities to liquefy it for transport and offering a significant safety risk. One potentially attractive option to get around this issue is to react hydrogen with CO_2 to give methanol,[3] a relatively stable liquid at room temperature which could be more easily handled using our existing petroleum infrastructure and used in unmodified combustion engines.[4] If the CO_2 required can someday be extracted directly from

the atmosphere this provides a convenient way of storing H_2, a carbon neutral fuel source and a green source of methanol for use in production of various chemicals (Figure 1). However, with the concentration of CO_2 in the atmosphere only at about 400 ppm this extraction may prove to be the biggest obstacle in the whole process, though capturing the CO_2 from industrial waste gas or volcanic sources where it is naturally more concentrated may offer a partial solution to the problem.[2,5,6]

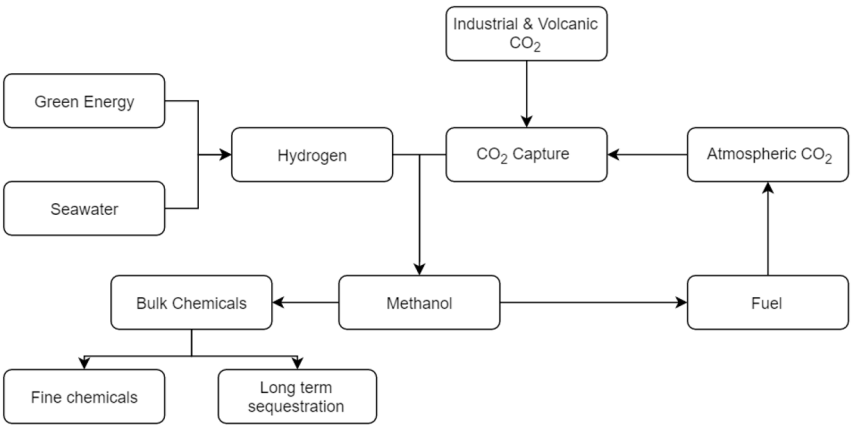


Figure 1

Carbon dioxide and hydrogen can react to give methanol according to the equation (all substances in their gaseous states):

$$1)\ CO_2 + 3H_2 \leftrightharpoons CH_3OH + H_2O\ ,\Delta H^{298} = -49\ kJ\ mol^{-1}, \Delta G^{298} = 3.77\ kJ\ mol^{-1} \quad [7]$$

The undesirable reverse water gas shift (RWGS) occurs simultaneously:

$$2)\ CO_2 + H_2 \leftrightharpoons CO + H_2O \qquad ,\Delta H^{298} = 41.2\ kJ\ mol^{-1}\ ,\Delta G^{298} = 28.62\ kJ\ mol^{-1} \quad [7]$$

Industrially methanol is not produced directly from CO_2 but rather from syngas, a mixture of CO, H_2 and CO_2 containing only a small percentage of CO_2.[1] The catalyst used is a mixture of Cu and ZnO supported on Al_2O_3 ($Cu/ZnO/Al_2O_3$).[8] It was long thought that the CO was the species being reduced but it is now accepted that the CO essentially serves as a source of CO_2 as well as a sink for water according to the above equations (eq. 1-2).[9–11] Catalysts used for the industrial production of MeOH with a syngas feed do not perform well under a pure CO_2/H_2 feed due to poisoning by water which builds up in the absence of CO and deactivation caused by sintering (combining of small particles into larger ones) of the Cu nanoparticles.[11,12] They also tend to give poor selectivity for methanol owing to a high activity towards the unwanted RWGS reaction (eq. 2).[11] Other side reactions to give methane, dimethyl ether and other hydrocarbons can also occur but are negligible with most of the catalysts discussed. Catalysts used in the production of MeOH from syngas can

be modified to give better results in a CO free feed and Cu based catalysts on metal oxide supports are the most common type for CO_2 hydrogenation both with and without CO present.[13] However even with modifications this type of catalyst experiences problems with deactivation and low selectivity and so development of an alternative is highly desirable.[11,14]

Thermodynamically the best yields of MeOH would be achieved at low temperatures and high pressures however in practise, due to the high kinetic stability of CO_2, temperatures in the region of 473-573 K, and pressures from 5.0 - 10.0 MPa are used.[8,15] High temperatures are the only way to get CO_2 to react to an appreciable extent but this poses problems for conventional catalysts. Increasing temperature accelerates their deactivation as well as favouring the production of CO and giving a rapidly falling space time yield (STY), measured in ($kg_{methanol}$)/($kg_{catalyst}$*hour). This creates demand for an alternative catalyst with improved selectivity and stability at increased temperatures.[13,16] Heterogenous catalysts tend to show less activity than homogenous ones however they are superior in terms of stability, separation, handling, and recyclability as well as price of reactor design and so they will be the focus of this review.[8,17,5]

There has been much research into improving current catalysts based on Cu combined with various metal oxide supports and these may prove the most effective option available, however they are already well researched and several new types of catalyst offer promising results.[13,18,19] Catalysts based on metals other than Cu and the use of novel support structures such as transition metal carbides (TMCs, compounds composed of carbon and a transition metal) and carbon nanotubes (CNTs) have given excellent yields of MeOH, in some cases performing better than traditional catalysts despite the relatively little refinement that they have undergone.[13] This review will discuss some of the most promising of these new catalysts based on metals other than copper as well as alternative support structures to the traditional metal oxides and whether they might offer a superior alternative to the standard industrial $Cu/ZnO/Al_2O_3$.

TMC Supported Catalysts

Transition metal carbides, particularly TiC and δ-MoC can be used as an alternative to metal oxides to support active metal clusters, as well as being catalytically active themselves.[20] Cu and Au clusters on both TiC and δ-MoC gave superior activity to a reference Cu/ZnO catalyst,[20,21] with Cu/δ-MoC giving the best performance, being 8-11 more active than the Cu/ZnO at 0.5 MPa over a temperature range of 500-600K.[21] However these results are all for very small metal clusters - only one atom thick on a metal carbide support (Figure 2) - so a great deal more research will be required to establish whether it can be made economically and whether they are stable over long reaction times.[22] TMCs also have the disadvantage of greatly accelerating the RWGS reaction, being more selective toward CO than MeOH.[20]

CO production is thought to occur at a different site to the hydrogenation though so it may prove possible to suppress this by intentional poisoning of these sites.[20]

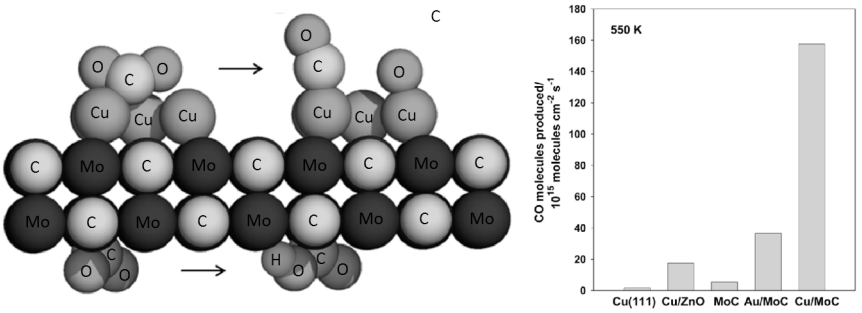


Figure 2: This shows the structure of the Cu/MoC catalyst with CO_2 reacting on both the metal clusters and on the carbide surface as well as a comparing activity to a reference Cu/ZnO catalyst.
Image adapted from Posada-Perez et. al,[21] with permission from Journal of the American Chemical Society.

CNT Supported Catalysts

The use of carbon nanotubes as a support structure and active component for Cu/ZrO_2 has also been investigated recently and shown promising results at 3 MPa (H_2:CO_2:N_2 = 69:23:8) and at temperatures of 493-515 K. The authors tested CNTs functionalised with basic nitrogen groups and acidic oxygen groups.[23] They found that nitrogen containing groups gave better performance. They attributed this to the nitrogen groups causing reduction of copper oxides and leading to smaller Cu particles and thus increased surface area as well as a reduced tendency to sinter, the opposite was observed for acidic oxygen functional groups as evidenced by XRD. They also found that the basic nitrogen was able to adsorb acidic CO_2 effectively by comparing the CO_2-TPD spectra of three different types of catalyst, one containing nitrogen groups and two without. Several different CNT structures were tested and performance varied quite significantly suggesting that even better results may emerge with continued research. The CNT supported catalysts tested offer activities and methanol STYs comparable to the reference Cu/ZnO and can likely be improved upon by the addition of different functional groups making them a promising option moving forward.[8] Interestingly, unlike metal oxides, the CNTs are hydrophobic and resist deactivation as water builds up. They also show improved stability over conventional catalysts, with yield staying roughly constant over 100 hours.[23]

Magnetically Active Catalysts

Iron is not generally used as a catalyst in the production of methanol from CO_2 as it shows poor selectivity for MeOH tending to produce other carbon containing products.[24,25] However, one interesting experiment conducted by Kiatphuengporn *et al*. involving Fe-Cu based catalysts supported on MCM-41 (a silicate/alumosilicate solid with pores of diameter 2-50 nm) may increase interest in the potential of ferromagnetic catalysts.[26] The authors found that the application of a 27.7 mT N-S magnetic field to the catalyst during reaction at 1.0 MPa and 493 K with a 1:3 ratio of CO_2 to H_2 increased MeOH selectivity 1.6 times. The application of various magnetic fields also led to an apparent decrease in activation energy of 1.12-1.15 times with the authors suggesting that this was due to the improved adsorbtion of reactants onto the magnetised surface. Fe was chosen over other ferromagnetic materials as it was the most active in CO_2 hydrogenation but the authors mention the possibility of using Co or Ni. The Fe-Cu catalyst tested was not particularly effective, performing poorly in comparison to other catalysts even under the applied magnetic field (though better results could be achieved by altering the pore size[27]). However, the consideration of magnetic properties in designing new catalysts for this reaction seems to be a very new idea and could potentially give much improved results in conjunction with more optimised magnetically active catalysts.

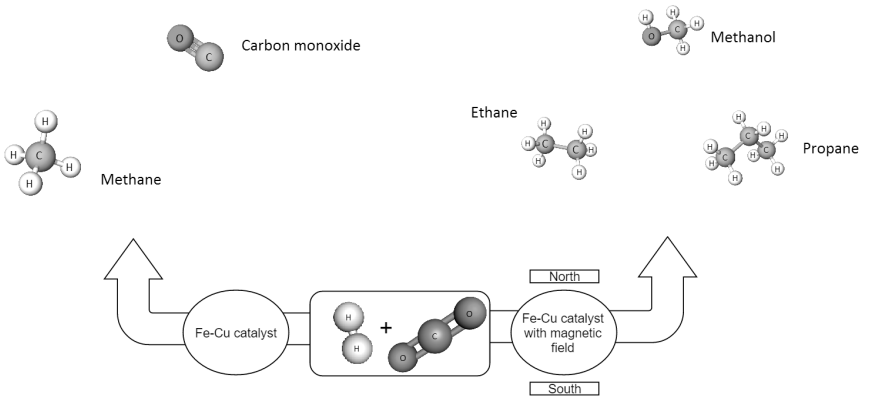


Figure 3: This diagram shows the products of CO_2 hydrogenation both with and without an applied magnetic field. Note the preference for hydrocarbons and methanol when a suitable magnetic field is applied.
Image adapted from Kiatpheungporn et. al,[26] with permission from Chemical Engineering Journal.

Non-copper Based Catalysts

Studt *et. al* produced a theoretical activity curve relative to pure Cu(211) for a number of different metals and alloys based on their binding strength for key

intermediate species in the reaction.[3] The model is somewhat simplified with DFT analysis performed based on the binding strength of eight potentially rate limiting key intermediates and without consideration of selectivity. It nevertheless agrees well with experimental evidence in most cases, placing Cu/Zn near the ideal binding strength and suggesting the success of several Ni-Ga compounds; though the most active such compound as suggested by the computational study proves to be ineffective due to rapid poisoning. It is worth noting that although Cu is the most active pure metal catalysts tested (which were Re, Rh, Ru, Fe, Co, Ni, Pd, Pt, Ag, Au, Cu),[28] it performs very poorly on its own.[15] Pd was found to be the next best pure metal to Cu under atmospheric conditions, which goes some way towards explaining why it is the next most commonly used metal in catalysts for this reaction.[9] As with copper it can be combined with other elements (Particularly ZnO as well as Gallium and its oxides)[29] to bring it closer to the ideal binding energy for the intermediates in CO_2 hydrogenation, though specific data was not available. Like Cu/ZnO, Pd/ZnO can be combined with a number of metal oxides to give an effective catalyst. It gives better selectivity than Cu/ZnO though this is unlikely to offset the much higher cost.[9,30] Pd/Zn/CNTs show very good methanol selectivity, however, with an activity much lower than that of conventional catalysts,[29] and less than the analogous Cu/Zn/CNT catalysts they are of limited interest in their current form.[8]

Based on results from Fujitani *et al*., Kong *et al*. suggested that a Ga-Pd alloy should be active towards MeOH production and this is indeed the case.[30,31] Li et. al tested a number of alloys and found Pd_2Ga to be the most promising.[32] The activity of a $GaPd_2/Ga_2O_3$ catalyst was five times higher than of $Cu/ZnO/Al_2O_3$ at 250C and 3.0 MPa (H_2:CO_2=3:1) and the production of undesirable CO was significantly lower.[32] However this activity is normalised by active metal content, by weight the Palladium catalyst is actually significantly worse.[32] This catalyst also had much greater stability, showing no loss of activity over 20 hours of use at 200 ^oC as compared to a 25% reduction in activity for the industrial one under the same conditions.[33] The same alloy with a high surface area SiO_2 (214 m^2/g) support shows increased activity and twice the methanol selectivity in the much lower temperature range of 478-483 K and at atmospheric pressure.[33] Selectivity with Pd is also more or less constant with temperature (50-60%) whereas with Cu catalysts methanol selectivity falls sharply at high temperature falling from 93% to 35% over a 60 degree range.[32] This higher selectivity at elevated temperatures is desirable as well as the increased stability, however the cost of Palladium compared to Cu may prove prohibitive in developing catalysts for widespread use based on it unless it proves to greatly outstrips Cu in terms of activity.

Nickel, like Cu and Pd, performs poorly in CO_2 hydrogenation however its alloys can be highly active. Several Ni-Ga compounds tested by Studt et. al. show much promise as alternative catalysts having activity better than traditional catalysts and a much lower rate associated with RWGS, especially at high temperatures.[3] While Cu/Zn is very close to the optimal binding energy value it has a relatively low density of catalytically active sites per gram of catalyst.[15] So while the activity

per site may be nearly ideal a material with a similar activity and a higher density of active sites is desirable. In traditional catalysts Cu is not itself highly active, showing good activity only in places where it is promoted by Zn. According to Studt *et. al* a higher density of sites may be easier to achieve with these intermetallic Ni/Ga catalysts since they do not rely on this promotion effect. Though a direct comparison of active site density was not made with the reference catalyst and they acknowledge that this is likely a factor in the difference in activity.[3] In terms of stability Ni-Ga alloys on a SiO_2 support are quite promising, activity for methanol production falls off rather slowly with temperature while activity towards the RWGS reaction falls much more sharply; the authors suggest distinct catalytic sites for both reactions with those mediating RWGS being poisoned by CO as temperature rises, unlike with Cu/Zn where both reactions take place at the same site.[3] The catalyst maintains 90% activity after 60 hours of reaction and can also be fully regenerated by passing H_2 at 623 K through it for two hours.

In_2O_3/ZrO_2 was compared to $Cu/ZnO/Al_2O_3$ by Martin *et al* under current standard industrial conditions (T = 473-573 K, P= 1.0 - 5.0 MPa, 1:4 ratio of CO_2 to H_2) and found to be greatly superior in terms of activity.11 This catalyst was found to be almost 100% selective over a range of temperatures (whereas the $Cu/ZnO/Al_2O_3$ catalyst peaked at 47% selectivity due to the RWGS reaction) with the STY increasing with temperature in contrast to the $Cu/ZnO/Al_2O_3$. It had the added advantage of being only moderately affected by H_2O poisoning, with its activity falling by just 8% after 400 hours and selectivity remaining constant at 99.8 %. Under the same conditions the activity of $Cu/ZnO/Al_2O_3$ fell by 44% in just the first 100 hours.[11] This catalyst has the advantage over many of the other novel catalysts mentioned of having a relatively simple structure making it easier to prepare for industrial scale operations, though the cost of Indium may prove prohibitive.

Conclusions

Some of the most promising alternative catalysts to traditional Cu/ZnO have been discussed in this review. Several catalysts proved superior to Cu/ZnO in terms of selectivity, stability and yield despite undergoing relatively little optimisation. However, while Cu/ZnO catalysts have proven successful on an industrial scale these newer catalysts have not and producing them in large amounts may prove impractical. In those that use expensive transition metals like Indium or exotic support structures like CNTs, cost may prove the limiting factor, though this may not be an issue if they show sufficient recycleability. Despite this, early results are very promising. While current catalysts are unlikely to undergo any drastic improvements newly developed ones offer much more scope for improvement given sufficient time and research.

Different catalyst preparation and reaction conditions make direct comparison of the catalysts impossible. However, given the early stage of research such a

comparison would not be useful even if conditions were standardised. The number of new catalysts being studied is itself very promising. If CO_2 can be extracted from the atmosphere the move to a methanol economy appears to be a very attractive option moving forward. It simultaneously gives a near inexhaustible source of fuel while offering a viable carbon source for a huge range of industrial chemicals with greatly reduced environmental impact compared to current sources. A highly efficient catalyst will be needed if this reaction is to be viable on a large scale but given the already promising results in several completely different types of catalyst this seems likely in the near future.

References

1. Olah, G. A.; Goeppert, A.; Prakash, G. K. S. Beyond Oil and Gas: The Methanol Economy: Second Edition; 2009.

2. Xu; Moulijn, J. a. Energy & Fuels 1996, 10 (2), 305–325.

3. Studt, F.; Sharafutdinov, I.; Abild-Pedersen, F.; Elkjær, C. F.; Hummelshøj, J. S.; Dahl, S.; Chorkendorff, I.; Nørskov, J. K. *Nat. Chem.* 2014, 6 (4), 320–324.

4. Ganesh, I. Renew. Sustain. *Energy Rev.* 2014, 31, 221–257.

5. Kothandaraman, J.; Goeppert, A.; Czaun, M.; Olah, G. A.; Prakash, G. K. S. *J. Am. Chem. Soc.* 2016, 138 (3), 778–781.

6. Kauw, M.; Benders, R. M. J.; Visser, C. *Energy 2015*, 90, 208–217.

7. James, A. M.; Lord, M. P. Macmillan's chemical and physical data; *Macmillan: London* [etc.], 1992.

8. Wang, W.; Wang, S.; Ma, X.; Gong, *J. Chem. Soc. Rev.* 2011, 40 (7), 3703.

9. Ma, J.; Sun, N.; Zhang, X.; Zhao, N.; Xiao, F.; Wei, W.; Sun, Y. *Catal. Today* 2009, 148 (3–4), 221–231.

10. An, X.; Li, J.; Zuo, Y.; Zhang, Q.; Wang, D.; Wang, *J. Catal. Letters* 2007, 118 (3–4), 264–269.

11. Martin, O.; Martin, A. J.; Mondelli, C.; Mitchell, S.; Segawa, T. F.; Hauert, R.; Drouilly, C.; Curulla-FerrMartin, O, D.; Pérez-Ramirez, J. Indium Oxide as a Superior Catalyst for Methanol Synthesis by CO2 Hydrogenation *Angew. Chemie - Int. Ed.* 2016, 55 (21), 6261–6265.

12. Samei, E.; Taghizadeh, M.; Bahmani, M. Fuel Process. *Technol.* 2012, 96, 128–133.

13. Porosoff, M. D.; Yan, B.; Chen, J. G. Energy *Environ. Sci.* 2015, 2 (2), 303.

14. Sun, J. T.; Metcalfe, I. S.; Sahibzada, *M. Ind. Eng. Chem. Res.* 1999, 38 (10), 3868–3872.

15. Behrens, M.; Studt, F.; Kasatkin, I.; Kühl, S.; Hävecker, M.; Abild-Pedersen, F.; Zander, S.; Girgsdies, F.; Kurr, P.; Kniep, B.-L.; Tovar, M.; Fischer, R. W.; Nørskov, J. K.; Schlögl, R.; Spencer, M. S.; Kasatkin, I.; Kurr, P.; Kniep, B.; Trunschke, A.; Schlögl, R.; Behrens, M.; Kurtz, M.; Wilmer, H.; Genger, T.; Hinrichsen, O.; Muhler, M.; Kurtz, M.; Yoshihara, J.; Campbell, C. T.; Rasmussen, P. B.; Szanyi, J.; Goodman, D. W.; Burch, R.; Golunski, S. E.; Spencer, M. S.; Kanai, Y.; Watanabe, T.; Fujitani, T.; Uchijima, T.; Nakamura, J.; Klier, K.; Ponec, V.; Jansen, W. P. A.; Frost, J. C.; Nakamura, J.; Choi, Y.; Fujitani, T.; Waugh, K. C.; Hansen, P. L.; Günter, M. M.; Chinchen, G. C.; Denny, P. J.; Jennings, J. R.; Spencer, M. S.; Waugh, K. C.; Grabow, L. C.; Mavrikakis, M.; Hu, Z.-M.; Takahashi, K.; Nakatsuji, H.; Yang, Y.; Evans, J.; Rodriguez, J. A.; White, M. G.; Liu, P.; Morikawa, Y.; Iwata, K.; Terakura, K.; Topsøe, N.-Y.; Topsøe, H.; d'Alnoncourt, R. N.; Grunwaldt, J. D.; Molenbroek, A. M.; Topsoe, N. Y.; Topsoe, H.; Clausen, B. S.; Słoczynski, J.; Fisher, I. A.; Woo, H. C.; Bell, A. T.; Kniep, B. L.; Behrens, M.; Behrens,

M.; Kasatkin, I.; Kühl, S.; Weinberg, G.; Chinchen, G. C.; Hay, C. M.; Vanderwell, H. D.; Waugh, K. C.; Hinrichsen, O.; Genger, T.; Muhler, M.; Bahn, S. R.; Jacobsen, K. W.; Laasonen, K.; Pasquarello, A.; Car, R.; Lee, C.; Vanderbilt, D.; Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D.; Vanderbilt, D.; Hammer, B.; Hansen, L. B.; Nørskov, J. K.; Monkhorst, H. J.; Pack, J. D.; Kresse, G.; Furthmuller, J.; Bengtsson, L.; Peterson, A. A.; Abild-Pedersen, F.; Studt, F.; Rossmeisl, J.; Nørskov, J. K. *Science* 2012, 336 (6083), 893–897.

16. Rodriguez, J. A.; Liu, P.; Stacchiola, D. J.; Senanayake, S. D.; White, M. G.; Chen, J. G. *ACS Catal.* 2015, 5 (11), 6696–6706.

17. Dai, W. L.; Luo, S. L.; Yin, S. F.; Au, C. T. Applied Catalysis A: *General.* 2009, pp 2–12.

18. Graciani, J.; Mudiyanselage, K.; Xu, F.; Baber, A. E.; Evans, J.; Senanayake, S. D.; Stacchiola, D. J.; Liu, P.; Hrbek, J.; Fernandez Sanz, J.; Rodriguez, J. A. *Science* (80-.). 2014, 345 (6196), 546–550.

19. Zhang, L.; Zhang, Y.; Chen, S. Appl. *Catal. A Gen.* 2012, 415–416, 118–123.

20. Posada-Pérez, S.; Viñes, F.; Rodriguez, J. a.; Illas, F. Top. *Catal.* 2014, 58, 159–173.

21. Posada-Pérez, S.; Ramirez, P. J.; Evans, J.; Villes, F.; Liu, P.; Illas, F.; Rodriguez, J. A. *J. Am. Chem. Soc.* 2016, 138 (26), 8269–8278.

22. Vida1. Vidal AB, Feria L, Evans J, Takahashi Y, Liu P, Nakamura K, *et al*. CO2 activation and methanol synthesis on novel Au/TiC and Cu/TiC catalysts. J Phys Chem Lett. 2012;3(16):2275–80. l, A. B.; Feria, L.; Evans, J.; Takahashi, Y.; Liu, P.; Nakamura, K.; Illas, F.; Rodriguez, J. A. J. Phys. *Chem. Lett.* 2012, 3 (16), 2275–2280.

23. Wang, G.; Chen, L.; Sun, Y.; Wu, J.; Fu, M.; Ye, *D. Rsc Adv.* 2015, 5 (56), 45320–45330.

24. Yan, S. R.; Jun, K. W.; Hong, J. S.; Choi, M. J.; Lee, K. *W. Appl. Catal. A Gen*. 2000, 194, 63–70.

25. Riedel, T.; Schaub, G.; Jun, K.-W.; Lee, K.-*W. Ind. Eng. Chem. Res*. 2001, 40 (5), 1355–1363.

26. Kiatphuengporn, S.; Jantaratana, P.; Limtrakul, J.; Chareonpanich, *M. Chem. Eng. J.* 2016, 306, 866–875.

27. Kiatphuengporn, S.; Chareonpanich, M.; Limtrakul, *J. Chem. Eng. J*. 2014, 240, 527–533.

28. Studt, F.; Abild-Pedersen, F.; Norskov, J. K.; Dahl, S.; Sharafutdinov, I.; Elkjaer, *C. F. Google Patents* 2012.

29. Liang, X. L.; Dong, X.; Lin, G. D.; Zhang, H. *Bin. Appl. Catal. B Environ*. 2009, 88 (3–4), 315–322.

30. Fujitani, T.; Saito, M.; Kanai, Y.; Watanabe, T.; Nakamura, J.; Uchijima, *T. Appl. Catal.* A, Gen. 1995, 125 (2).

31. Kong, H.; Li, H. Y.; Lin, G. D.; Zhang, H. *Bin. Catal. Letters* 2011, 141 (6), 886–894.

32. Li, L.; Zhang, B.; Kunkes, E.; Föttinger, K.; Armbrüster, M.; Su, D. S.; Wei, W.; Schlögl, R.; Behrens, M. *ChemCatChem* 2012, 4 (11), 1764–1775.

33. Fiordaliso, E. M.; Sharafutdinov, I.; Carvalho, H. W. P.; Grunwaldt, J. D.; Hansen, T. W.; Chorkendorff, I.; Wagner, J. B.; Damsgaard, C. D. *ACS Catal*. 2015, 5 (10), 5827–5836.

TS
SR

Prolonging the Lifespan of Lithium-Ion Batteries with Nanostructured Silicon Anodes

Dáire Brady
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

Since their commercial availability in 1991, lithium-ion batteries (LiB) have proven to be the most successful and widely used source of power in the market of consumer electronics. Their light weight and compact size have accelerated miniaturization, while their higher energy density and longer lifespan have wide ranging applications in mobile electronics, photovoltaics, biological sensors and much more. Research continues in this field 26 years later, with emerging tech markets (electric vehicles, grid energy storage, etc.) heralding in an increasingly electric future. LiB production increases every year in response to this, however there is an insatiable demand for increasingly efficient, cost-effective and longer lasting batteries. LiBs with nanostructured silicon anodes (nanoSi/LiB) pose promising solutions to this challenge, with marked improvements in power density over traditional anodes, however pulverisation and loss of charge remain challenges to be overcome. This review explores the different strategies and techniques researchers are using to tackle this problem, with commentary on the potential of nanoSi/LiBs for future demands and applications being presented.

Introduction

Lithium is the lightest and one of the most electropositive metal on the periodic table, making it a very mobile, energy dense charge carrier in an electrolyte. The past two decades have seen staggering progress in its research in LiBs, whose potential and scientific interest has yet to be exhausted. Having reach ubiquity in all modern electronics they have greater energy densities, efficiencies and lighter weights than most secondary batteries. A mature technology on the small scale, efforts are being made to upscale their potential for use in larger energy storage

projects. Companies like Tesla Motors are at the forefront of this revolution, who are working with other utilities to build the world's largest LiB grid network that will supply energy to the Los Angeles area. Having gone online from January 2017, construction is expected to be complete by 2021.[1] Meanwhile, construction is ongoing for the largest building in the world - "Gigafactory" - which by 2020 will produce more LiBs per year than were globally produced in 2013 for use in their electric vehicles.[2] Home battery storage has also become an emerging technology in recent years, which can provide considerable economic and environmental advantages when coupled with renewable technology.[3] With Tesla and other companies showing initial success in these projects, it is foreseeable that many enterprises will follow suit. It is therefore in the public interest that research continues to meet these accelerating demands.

Despite the promising investment in LiBs however, there are some drawbacks that current science research is working to resolve. One particular issue is the poor capacity and lifetimes of conventional LiBs which typically stem from the anode. Commercially produced since 1997, graphite has been a widely successful anode for LiBs due to its low cost, high structural integrity, electron affinity and electronic conductivity.[4] However, the advancement of technology has revealed within it many shortcomings that make it less attractive to continued development. The main hindrance of graphite lies in its electrochemical reaction with Li^+, in which it takes six carbon atoms to bond to each lithium ion, giving a maximum theoretical charge capacity of only 372 mAh g^{-1}.[5,6] Batteries of this capacity are fast becoming obsolete in society as the crux behind large scale developments like grid-energy storage and electric vehicles (EV) stems from their short lifespan, high cost and heavy, bulky form.

Of all the materials to succeed graphite, silicon proves the most promising. At 4200 mAh g^{-1},[7] it has ten times the theoretical capacity of graphite, with each silicon atom bonding to up to 4.4 Li^+ ions. Initial researched confirmed this capacity increase, however they were shown to fade by 80% after as little as 10 cycles.[8] This surprising result arose from the expansion of silicon by ~400% in the lithiated state,[6] inducing considerable stress and strain on the bulk which led to cracking, pulverization, delamination from the current collector and irreparable capacity loss. Research has since then focused on tackling this issue by restructuring the anode on the nanoscale. NanoSi displays a number of advantages over bulk silicon. It displays the same exceptional capacities and conductances of bulk silicon but doesn't degrade as quickly from volume changes. A larger surface area in contact with the electrolyte also increases the diffusion of Li^+ into the anode allowing faster charging times.

The review presents the current state of research in these nanoSi anodes, with analyses into the structures that have yielded the most scientific interest, as well as the efforts being made to increase their scalability and cost-effectiveness for industry.

The Promises of Pristine Silicon Nanowires (SiNWs)

Chan *et al*. presented a major milestone in the field of energy storage with their use of SiNWs,[5] reaffirming that miniaturization to the nanoscale was a step in the right direction.[9,10] As with all forms of nanoSi, SiNWs displayed a number of advantages over bulk silicon, such as faster charging times, greater capacities and greater resilience to pulverization from the decreased stress of smaller structures[11] (Figure 1). The first charge capacity agreed with the theoretical capacity of Si at 4277 mAh g^{-1} when cycled at the C/20 (i.e. 1 charge/discharge per 20 hours) and were also shown to maintain high capacities at higher rates, stabilising to 2100 mAh g^{-1} after 20 cycles at 1C, which is 5 times greater than graphite. The experiment made Si an increasingly viable anode material for commercial use, having shown marked improvements in capacity retention. Furthermore, these SiNWs were grown directly on a stainless steel current collector to eliminate the need for non-conducting binding agents[12] using the Vapour-Liquid-Solid Method (VLS) method. By itself, the VLS method has great potential for scalability to larger scale production, due to its simplicity and relatively few steps,[13] but the high temperatures required by the Chemical Vapour Deposition (CVD) process restrict this production to the lab scale. VLS remains to be the most studied and used bottom-up method of NW growth however, and efforts are being made to find more cost-effective modes of NW growth, such as plasma-assisted VLS growth[13] which significantly reduces work temperatures. Interestingly, these developed methods have paved the way for the use of conductive flexible substrates on which to grow NWs, which in itself has obvious implications for flexible electronics and energy storage.[14]

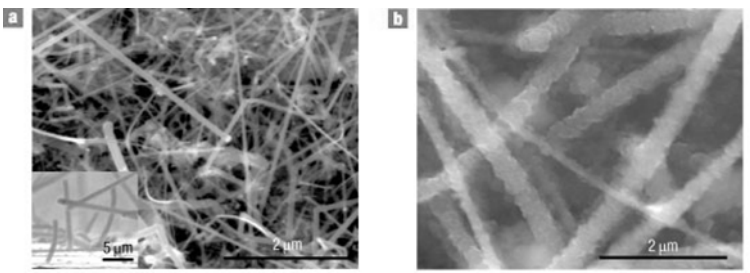


Figure 1: (Chan *et al*.) SEM images of pristine SiNWs (A) before and (B) after cycling of the battery. The images show the resilience of SiNWs to pulverisation during the lithiation process.

Double-Walled Silicon Nanotubes (DWSiNTs)

DWSiNTs have led to some of the most eye-opening results in recent years, having been shown to maintain a charge capacity of 85% after an astonishing 6000 cycles. To date this has not been demonstrated with any other form of nanoSi cycled under the same conditions. This was demonstrated by Wu *et al*.[15] in 2012, who adopted a very creative approach which almost entirely eliminate anode pulverisation.

The main focus of the group was to prevent the formation of an unstable Solid Electrolyte Interphase (SEI), which forms on the anode from electrolyte side reactions during cycling. The SEI layer completely covers the Si anode after lithiation when its volume has reached its maximum, and exposes more Si to the electrolyte by breaking up during delithiation when the anode contracts to its original size. When lithiation starts again, more SEI forms on these exposed surfaces, and this process continues as the SEI gets thicker (Figure 2). This poses many problems to the performance of the battery such as (i) Reduction in charge available due to Li^+ consumption; (ii) Weakening of contact between anode and current collector due to the electronically insulating SEI and (iii) Mechanical stress and degradation of the anode.

The outer wall of the DWSiNT is composed of SiO_2, which doesn't expand during lithiation, thus limiting SEI formation to the first cycle. It also prevents the inner wall from coming into contact with the electrolyte but allows Li^+ to permeate through, giving it space to lithiate and delithiate without electrolytic disturbance.

At a rate of 10C, these nanoSi/LiBs retained an incredible 88% of its initial discharge capacity after 6000 cycles. The Coulombic Efficiency (CE - discharge capacity as a percentage of charge capacity) was also shown to be remarkably high (Figure 3(b)), with the average being 99.938% between the 2nd and 6000th cycle (The CE of the first cycle was 76% due to the consumption of lithium during SEI formation). The specific charge capacity was also exceptionally high, with capacities of up to 1780 mAh g^{-1} at C/5 and 600 mAh g^{-1} at 12 C (capacities taken over entire wait of DWSiNTs).

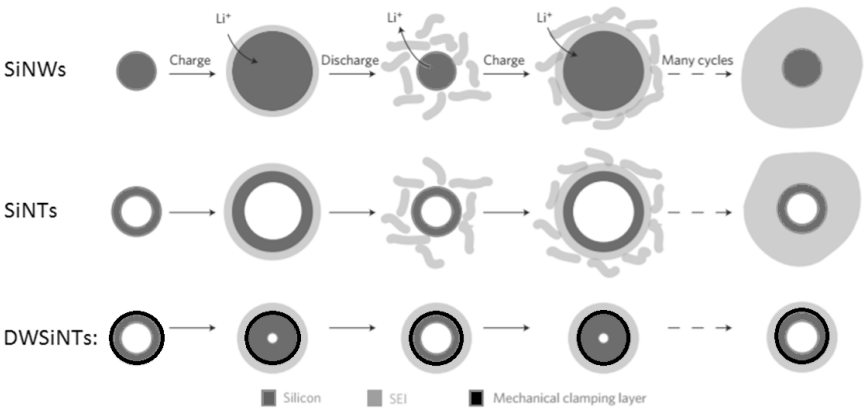


Figure 2: The formation of SEI around SiNWs SiNTs and DWSiNTs. Build up of SEI around SiNWs and SiNTs is a result of repeated expansion and contraction, whereas the outer wall of DWSiNTs maintains the same size and thus keeps the SEI layer thin.
Image adapted from Wu et al.

Despite these ground-breaking results however, DWSiNTs are not commercially used in nanoSi/LiBs There are a number of complex processes required in their production - of which CVD is one, making this experiment more of a proof-of-concept. As a result, lower-cost forms of nanoSi are more commonly considered in research, as their performance can be improved by structural modifications with other materials.

Coating of SiNWs

A 2014 paper by Kohandehghan *et al*. investigates the use of tin as a coating for SiNWs[16] (Sn/SiNWs). With a theoretical capacity of 996 mAh g^{-1}, tin is electrically conductive as well as structurally supportive to SiNWs, expanding by ~260% during lithiation. This generates a compressive stress on Si that favours longitudinal over radial expansion, thus reducing cracking. Having cycled Sn/SiNWs with various coating thicknesses, the optimum result obtained was from Sn/SiNWs coated with a 3 nm thick layer of tin. At 1865 $mAhg^{-1}$, it had almost twice the capacity of pristine SiNWs at 0.1 C (1046 mAh g^{-1}) and had a (CE) of 99% after 100 cycles. Wu *et al*.[17] achieved 90% capacity after an astonishing 5000 cycles by coating SiNPs in a conducting hydrogel polymer. Phosphoric acid groups in the polymer composite bind to SiO_2 on the anode surface, causing conformal coating of the anode and further cross-linking of monomers. Electrostatic interactions between the negative surface oxide and positive polyaniline also aided in creating a continuous

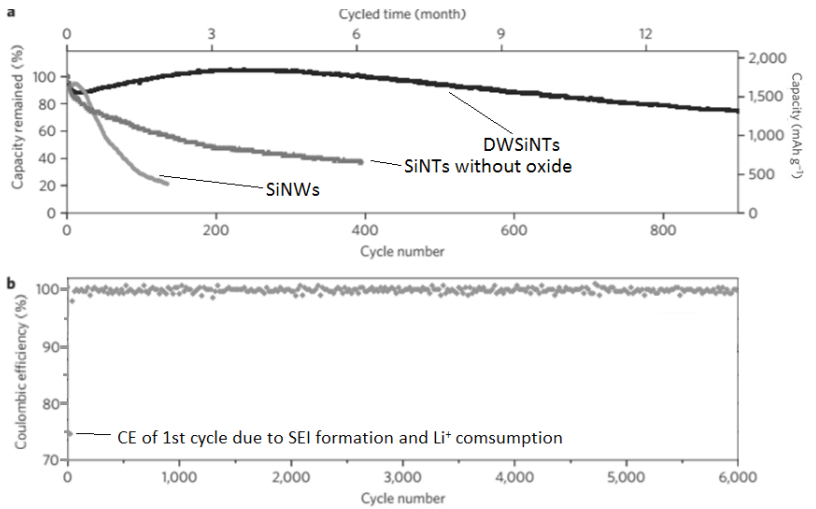


Figure 3:
(A) The capacity of DWSiNTs, SiNTs and SiNWs at C/5. The latter two were shown to degrade considerably faster than DWSiNTs. The relatively large decrease of DWSiNTs was possibly due to electrolyte leakage over the 12 months that the experiment was run.
(B) CE of DWSiNTs cycled at 12 C over 6000 cycles.
Image adapted from Wu et al.

electron and ion-conducting 3D network. The porosity of the network allowed ample space for SiNPs to swell into, and kept them trapped within the network should pulverisation occur. These findings have many promising implications, as the method of production is described as scalable and compatible with slurry-coating methods that are currently popular in industry. These papers successfully demonstrate the ability of a coating to provide structural and electrical support to nanoSi - an idea which is continuously seen in literature.[12,19,20,18]

Improving Cost and Scalability of NanoSi/LiB Production

As nanoSi/LiB technology becomes more mature methods are continuously being sought to make their production industrially compatible. Not only does this apply to LiBs, but also other fields such as the next generation of solar cells.[13] A paper released in 2016[21] investigates the use of a conductive polymer PEDOT:PSS as both binder and conductive additive to the SiNP anode. One of the main focuses of the paper was to bring cost-effective advances to the market, which was achieved by using the commonly used method of slurry casting (mixing active material, binder, solvent and other material into a paste). The polymer binder replaces traditional mixtures of conductive additives (carbon black, graphene and graphite) and inert polymer binders, thereby maximising spatial efficiency for conductive additives and increasing charge capacity. It also reduces capacity loss as the interface between it and the anode is considerably stronger than with previous conductive additives. Futhermore, samples of the anode where the PEDOT:PSS polymer was doped with up to 20 wt % formic acid (FA) were tested, as FA has been shown to improve conductivity in PEDOT:PSS,[22] and has self-healing properties SiNP electrodes.[23,24] This increased conductivity by up to 2 orders of magnitude. From the 2nd to 100th

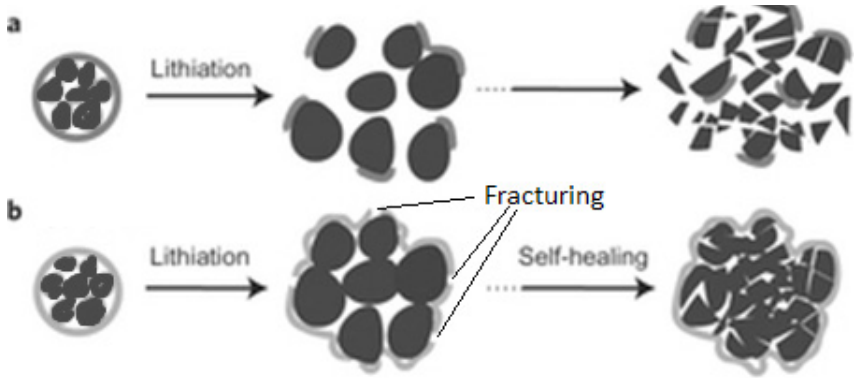


Figure 4:
(A) Degradation of SiMPs with traditional polymer binder compared to
(B) SHPs able to maintain structual integrity of anode with repeated cycling.
Image adapted from Wang et al.

cycle at 1 A g^{-1}, the capacity of 10% FA PEDOT:PSS/SiNP went from 2681 to 1950 mAh g^{-1}, which was a notable improvement over conventional LiBs with graphite anodes, which decreased in capacity from 2480 to 1510 mAh g^{-1}.

Bypassing Capacity Loss Issues in Silicon Microparticles (SiMPs)

Despite Chan *et al*'s commentary on the inefficacy of SiMPs,[5] Wang *et al*[23] found inspiration from nature by demonstrating that pulverisation and capacity loss can be overcome with the addition of self-healing polymer (SHP) binders. The experiment tested this on SiMPs 3-8μm in size and found that 90 cycles passed before the anode capacity reached 80% compared to <9 cycles for SiMPs with traditional binders and maintained a capacity of 98.5% at the C/10 rate. The SHPs in question were capable of stretching to 300% of their size compared to 2% for alignate binders such that if any SiMPs fractured during lithiation, they were still contained within the binder and maintained contact with other SiMPs (Figure 4). Despite undergoing some amount of fracturing, the SHPs completely or significantly repaired after several hours due to their tendency to spontaneously self-repair *via* H-bonding. While these improvements don't compete with results obtained from nanoSi, this experiment demonstrates the hidden potential of SiMPs in LiBs whose production methods possess significant cost and scalability improvements over nanoSi. Perhaps by combining these SHPs with a coupling electrolyte that aids self-healing, SiMPs could become a very competitive option for LiBs.

Conclusion

Research into LiBs has been ongoing since the 1970's, however the use of nanoSi has led to explosive growth in this field the past decade. Since Chan *et al*'s[5] work on SiNWs, virtually all forms of nanoSi have been researched, such as SiNPs[28,29] SiNTs,[30] DWSiNTs,[15] porous Si[31,32] and Si nanofibers.[33]

Performance enhancements of nanoSi/LiBs have also been achieved *via* structural modifications with other materials. The problem of pulverization can be minimised by protective coatings such as tin, or even circumvented by trapping pulverised nanoSi in a conductive polymer coating. Electrolyte composition can contribute to the anode by "synergistic coupling",[34] alleviating stress and strain and even the binder can contribute to its capacity.

As mentioned, DWSiNTs display the best lifetimes of any form of nanoSi under the same conditions, however their means of production are incredibly complex and unscalable. Now that research into nanoSi/LiBs has reached maturity, it is

important that further research be carried out to improve the cost and industrial compatibility of their production methods. SHP-coated SiMPs may also be a source of hidden potential should the research be developed, with cost a key advantage for competitiveness. Their self-healing chemistry is particularly interesting and could perhaps be coupled with other components of the battery to improve performance.

The vast scope of diversity in the research of LiBs holds a lot of promises for the future, however it's evident there are many challenges to be overcome. While capacities and lifetimes of nanoSi/LiBs are approaching those necessary for EVs and grid storage, the cost-effectiveness of production is still an issue limiting their access to the greater market. With decreasing costs, nanoSi/LiBs have great potential to contribute to the struggle against climate change. EVs and domestic energy storage are already starting to become economical and could eventually be adopted in developing nations, where rapid population growth and industrialization has seen increased strain on natural resources. This may be a long-term aspiration, however it's a necessary motive for present-day research.

It is difficult to predict the advances that will be made in nanoSi/LiBs over the next 5 years, as research shows no signs of subsiding. With the advent of larger scale technologies on the horizon, the breakthrough to bring us into the 20s will not be in futher enhancement, but in a means of making high performance LiBs economically viable and industrially compatible. Nevertheless, the vast scale of investment and engineering into this well-established technology is testament enough to its increasing influence in every aspect of society as a viable source of clean power for many years to come.

References

1. Fialka, J. *ClimateWire* 2016

2. Tesla Motors, *Blog* 2014, 44.

3. Naumann, M.; Karl, R. C.; Truong, C. N.; Jossen, A.; Hesse, H. C. Lithium-ion battery cost analysis in PV-household application. *Energy Procedia*. 2015; pp 37–47.

4. Billaud, J.; Bouville, F.; Magrini, T.; Villevieille, C.; Studart, *A. R. Nature Energy* 2016, 1, 1–6.

5. Chan, C. K.; Peng, H.; Liu, G.; McIlwrath, K.; Zhang, X. F.; Huggins, R. A.; Cui, Y. High-performance lithium battery anodes using silicon nanowires. *Ph.D. thesis*, 2008.

6. Rahman, M. A.; Song, G.; Bhatt, A. I.; Wong, Y. C.; Wen, C. Advanced *Functional Materials* 2015, 647–678.

7. Kim, J. S.; Choi, W.; Byun, D.; Lee, J. K. *Solid State Ionics* 2012, 212, 43–46.

8. Liu, X. H. *et al*. Nano Letters 2011, 11, 3312–3318, PMID: 21707052.

9. Green, M.; Fielder, E.; Scrosati, B.; Wachtler, M.; Moreno, J. S. *Electrochemical and Solid State Letters* 2003, 6, A75.

10. Kasavajjula, U.; Wang, C.; Appleby, A. J. Nano- and bulk-silicon-based insertion anodes for lithium-ion secondary cells. 2007.

11. Liu, X.; Long, Y. Z.; Liao, L.; Duan, X.; Fan, Z. Large-scale integration of semiconductor nanowires for high-performance flexible electronics. 2012.

12. Memarzadeh, E. L.; Kalisvaart, W. P.; Kohandehghan, A.; Zahiri, B.; Holt, C. M. B.; Mitlin, D. *Journal of Materials Chemistry* 2012, 22, 6655.

13. Misra, S.; Yu, L.; Chen, W.; Foldyna, M.; Cabarrocas, P. R. I. *Journal of Physics D: Applied Physics* 2014, 47, 393001.

14. Xie, X.; Zeng, X.; Yang, P.; Wang, C.; Wang, Q. *Journal of Crystal Growth* 2012, 347, 7–10.

15. Wu, H.; Chan, G.; Choi, J. W.; Ryu, I.; Yao, Y.; McDowell, M. T.; Lee, S. W.; Jackson, A.; Yang, Y.; Hu, L.; Cui, Y. *Nature Nanotechnology* 2012, 7, 310–315.

16. Kohandehghan, A.; Cui, K.; Kupsta, M.; Memarzadeh, E.; Kalisvaart, P.; Mitlin, D. *Journal of Materials Chemistry* A 2014, 2, 11261–11279.
17. Wu, H.; Yu, G.; Pan, L.; Liu, N.; McDowell, M. T.; Bao, Z.; Cui, Y. *Nature communications* 2013, 4, 1943.

18. Wang, X.; Li, G.; Seo, M. H.; Lui, G.; Hassan, F. M.; Feng, K.; Xiao, X.; Chen, Z. W. *ACS Applied Materials and Interfaces* 2016

19. Tang, D.; Yi, R.; Gordin, M. L.; Melnyk, M.; Dai, F.; Chen, S.; Song, J.; Wang, D. J. *Mater. Chem*. A 2014, 2, 10375–10378.

20. Kohandehghan, A.; Kalisvaart, P.; Kupsta, M.; Zahiri, B.; Amirkhiz, B. S.; Li, Z.; Memarzadeh, E. L.; Bendersky, L. a.; Mitlin, *D. Journal of Materials Chemistry* A 2013, 1, 1600.

21. Higgins, T. M.; Park, S. H.; King, P. J.; Zhang, C.; McEvoy, N.; Berner, N. C.; Daly, D.; Shmeliov, A.; Khan, U.; Duesberg, G.; Nicolosi, V.; Coleman, J. N. *ACS Nan*o 2016, 10, 3702–3713.

22. McCarthy, J.; Hanley, C.; Brennan, L.; Lambertini, V.; Gun'ko, Y. J. *Mater. Chem.* C 2014, 2, 764–770.

23. Wang, C.; Wu, H.; Chen, Z.; McDowell, M. T.; Cui, Y.; Bao, Z. *Nature Chemistry* 2013, 5, 1042–8.

24. Kwon, T.-w.; Jeong, Y. K.; Lee, I.; Kim, T.-S.; Choi, J. W.; Coskun, *A. Advanced Materials* 2014, 26, 7979–7985.

25. Yae, S.; Morii, Y.; Fukumuro, N.; Matsuda, H. *Nanoscale Research Letters* 2012, 7, 352.

26. Hu, Y.; Peng, K.-Q.; Liu, L.; Qiao, Z.; Huang, X.; Wu, X.-L.; Meng, X.-M.; Lee, S.-T. *Scientific Reports* 2014, 4, 3667.

27. Ge, M.; Rong, J.; Fang, X.; Zhou, C. *Nano Letters* 2012, 12, 2318–2323, PMID: 22486769.

28. Sun, L.; Su, T.; Xu, L.; Du, H.-B. Physical Chemistry Chemical Physics 2016, 18, 1521–1525. [29] Pan, L.; Wang, H.; Gao, D.; Chen, S.; Tan, L.; Li, L. *Chemical communications (Cambridge, England)* 2014, 50, 5878–80.

30. Wen, Z.; Lu, G.; Mao, S.; Kim, H.; Cui, S.; Yu, K.; Huang, X.; Hurley, P. T.; Mao, O.; Chen, J. *Electrochemistry Communications* 2013, 29, 67–70.

31. Zhu, J.; Gladden, C.; Liu, N.; Cui, Y.; Zhang, X. *Physical chemistry chemical physics* : PCCP 2013, 15, 440–3.

32. Li, X.; Gu, M.; Hu, S.; Kennard, R.; Yan, P.; Chen, X.; Wang, C.; Sailor, M. J.; Zhang, J.-G.; Liu, *J. Nature Communications* 2014, 5, 1–7.

33. Cho, D.; Kim, M.; Hwang, J.; Park, J. H.; Joo, Y. L.; Jeong, Y. *Nanoscale research letters* 2015, 10, 424.

34. Park, S.; Bok, T.; Cho, S.-J.; Choi, S.; Choi, K.-H.; Park, H.; Lee, S.-Y. *RSC Adv*. 2016, 6, 6960–6966.

TS
SR

Redox Flow Batteries: A Grid Energy Storage Solution?

Tim Ryan
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

Flow Battery research has accelerated rapidly in the last few years as researchers attempt to find a long term energy storage solution for our planet. Flow batteries are promising due to their relatively long cycle life, low cost per unit storage and high working efficiency as compared to the solid state technologies (eg. LiIon) which dominate the market today. Vanadium flow batteries are currently the most mature and market ready technology, with a significant amount of effort currently dedicated to their improvement toward a target cost of $120/kWh. While this effort is ongoing, there is also significant promise being shown in alternative flow battery technologies such as lead, lithium and zinc-bromine. The use of nanomaterials and nanoscale techniques is showing particular promise. Here, we discuss recent developments in flow battery research, keeping energy efficiency and cost in mind. Particular attention is given to Vanadium and Lead flow technologies, which appear particularly promising for use in large scale grid energy storage. Multiple other flow technologies are discussed, including the emerging fields of organic and non-aqueous flow batteries, which have huge scope for future development.

Introduction

In 2014, Global energy production was the equivalent of nearly 160,000 TWh.[1a] Of this, 81.1% was provided by fossil fuels.[1a] By 2040, we are expected to require the equivalent of a further 50,000 TWh in energy.[1b,2] As such, energy production must increase. However, due to the threat of Climate Change and the Paris agreement, which aims to keep warming "well below 2°C",[3] we require a massive reduction in current levels of CO_2 emission.[4] Thus, new energy must come from non-carbon emitting sources. Renewable energies, such as Wind, Solar and Tidal are considered

to be a large part of the solution, with each recently experiencing massive growth.[2,5] However these technologies all have the inherent problem of intermittence. Unlike fossil fuel power stations, the amount and timing of energy production cannot be controlled and can be hard to predict.[2,5] This presents a challenge in balancing demand to ensure power network stability and reliability. As renewable energies become the main source of power, it is clear that Electric energy storage (EES) in some form will be required.[5,6] Currently we have the capacity worldwide to store only 1% of the energy we consume.[2] In the coming years, our ability to store energy will need to dramatically increase.[2,7] For economic viability, a cost target of 120$/kWh (given a 10 year lifespan) is required.[8] As no technology is currently at this level, multiple technologies are being explored, from massive iron flywheels to compressed air storage.[9] However Electrochemical Energy Storage (EES) systems are considered the most likely to be able to provide the required capacity, efficiency, power and response time at the required cost.[9]

Multiple types of EES systems exist. A number are compared in Table 1 (adapted from[10,11]). For large scale grid energy storage, technologies must meet multiple targets. The technology must be capable of scaling to MW levels of energy storage with high energy efficiency. Energy efficiency is quoted in two ways;

Technology	Typical Power (MW)	Discharge Time	Storage Capacity Cost ($/kWh)	Life Time (cycle/years)	Efficiency (%)	Energy Density (Wh/kg)	Drawbacks
Supercapacitors	0.25	<1 min	500-3000	500,000/20	>90	4.5-10	Explosion hazard, low energy density, cost
Regenerative Fuel Cells with Hydrogen Storage	10	>5 h		13	40-50		Low-density storage, high cost, safety
Lead-acid Batteries	0.5-20	3-5 h	65-120	1000-1200/3-4	70-80	25-50	Low energy density, short lifetime, temperature sensitive
Li-ion Batteries		1-5 h	400-600	750-3000/6-8	80-90	70-220	Cost, safety, short lifetime, self-discharge, temperature sensitive
NAS Battery	0.25-1	6-8 h	360-500	2500-4500/6-12	87	103	Cost, high-temperature operation, safety
Flow Battery (VRB)	0.5-12	10 h	180-2500	500-2000/10	70-85	5-30	Low energy density

Table 1: Comparison of current Properties of Electrochemical Energy Storage Methods *Adapted from*[10,11]

Coulombic efficiency, the ratio of output charge to input charge and Energy efficiency; the ratio of output energy to input energy. For grid scale energy storage, energy efficiency of at least 80% is required. Cuolombic efficiency is expected to be higher again. Length of operating life is also important, with 10 years expected as a minimum.[12] Operating life is most affected by capacity fading, which occurs by multiple mechanisms. A discharge time in the scale of hours is advantageous, to make up for extended dips in energy production (eg. cloudy, non-windy day). Energy density is also important, however in stationary large scale storage, relatively low energy density (eg. 10 Wh/kg) is usually acceptable. Finally, the technology must be cost effective.[8]

Taking all of these demands into account, table 1 suggests that lead acid and flow batteries are most suitable for large scale grid energy storage. However, Lead acid batteries are seen to have too short a lifetime, and flow batteries too high a cost. When comparing the published literature, it is clear that lead-acid has been much heavier researched in the past,[13] which suggests that redox flow batteries, which have been relatively ignored until recently (see Fig. 2a) may have greater room for improvement. As such, this review will focus on the growing field of flow battery research which is showing great promise for use in large scale grid energy storage.[10,12]

Redox flow batteries (RFBs) operate differently from the more developed battery technologies (such as Lithium Ion, lead acid) as the electroactive materials move during operation.[10,14] Power and Energy are also uncoupled in RFBs which prevents the energy waste (causing a less than quoted working efficiency) due to excess power/ energy found in the more developed solid state electrode batteries.[10] RFB's also allow battery components to be easily replaced during and at end of life, extending operational lifetime without enormous additional costs.[15] A Schematic and basic operating principle of an RFB is shown in Figure 1.

Redox Flow Battery Technology

In RFBs, the energy capacity is controlled by the size of the electrolyte tanks (reservoirs) and power controlled by the size of the electrode/ cell stack.[11,15] Tank/ stack size is easily increased, meaning RFBs are capable of storing large amounts of power and energy (up to MW/MWh). Currently however, the high cost per kW and kWh, although still lower than many other battery technologies (Table.1) prevents large scale application.[10,11] Thus cost reduction is of high priority for current research in the area.

Redox Flow batteries have undergone slow but continuous development since their invention in the 1970's,[17] however the number of paper published on the topic has risen dramatically in recent years (see Fig. 2a).[12]

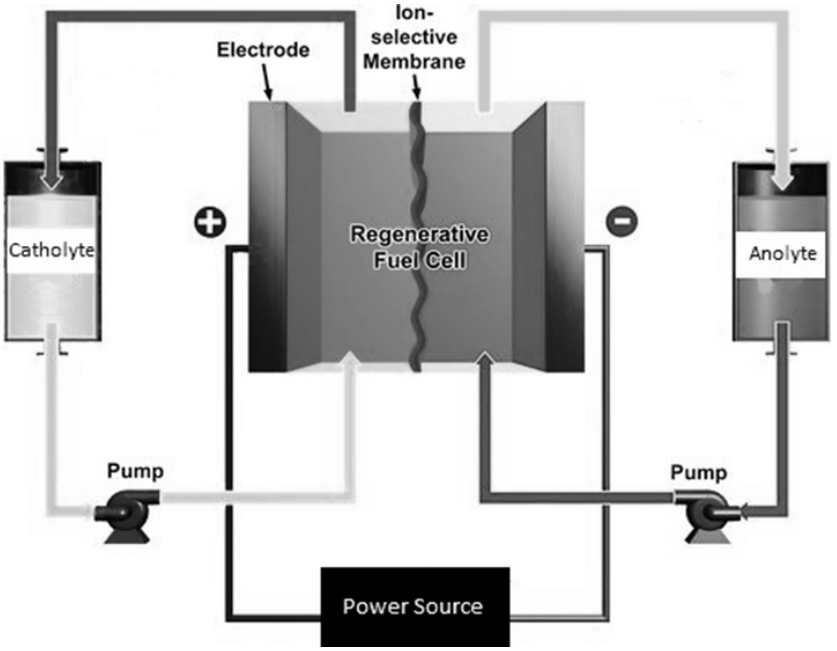


Figure 1: Basic schematic of a flow battery.[16] While charging, the input current (D.C) produces a negative and positive electrode. At the negative electrode (cathode), the anolyte (which is being pumped past) is reduced $V^{3+} \rightarrow V^{2+}$) , and at the positive electrode (anode), the catholyte is oxidised (e.g. $V^{4+} \rightarrow V^{5+}$). Positive ions (e.g. H^{+}) then cross the ion selective membrane (which prevents the electrolytes from mixing but allows the charge balancing ion through) to balance out the charge. The charged ions (in catholyte and anolyte) are then pumped back into the respective tanks and stored.[11,15]

The first RFB used an Fe^{3+}/Fe^{2+} halide solution in the positive half cell, and a Cr^{3+}/Cr^{2+} solution in the negative half cell, which quickly ran into the problem of cross contamination between the cells, resulting in dramatic capacity reductions.[15] In the 80's, systems of mixed electrolytes and Vanadium ions with greatly reduced cross contamination were developed, allowing much longer operational life.[10,18,19] Vanadium RFBs were a particularly promising technology, however at the time, the focus was on energy density, meaning the number of researchers working on them has been small until recently (Fig 2a).

Vanadium Redox Flow Batteries

A VRB takes advantage of the 4 oxidation states of Vanadium ions, with two redox couples separated by an ion selective membrane. This prevents the active species (vanadium ions) from crossing between anolyte and catholyte while allowing

charge to transfer in form of protons. The fact that vanadium is the active element on both sides makes this easier. Energy conversion (electric to chemical and vice versa) occurs through the half cell reactions illustrated below (1,2), where the forward reactions are the discharge reactions.[15]

$$\textit{Cathode side: } VO_2^+ + 2H^+ + e^- \leftrightharpoons VO^{2+} + 2H_2O \qquad (+1V) \qquad (1)$$

$$\textit{Anode side: } V^{2+} - e^- \leftrightharpoons V^{3+} \qquad (+0.26V) \qquad (2)$$

The overall electrochemical reaction gives a cell voltage of 1.26V at 25°C.[15,18] Sulfuric acid (as high as 4M) was traditionally used as the electrolyte in which the vanadium is dissolved, ensuring stability of the reaction.[10] However the solubility of the Vanadium species in the sulfuric acid limits the system energy density and operational range. Vanadium RFB is prone to precipitation of V_2O_5 at temperatures above 40°C in the catholyte which limits its operational range to 10 − 40°C with 1.7M Vanadium.[15]

Much research has focused on stabilising the catholyte in order to increase vanadium solubility and thus energy density. Additives such as ammonium oxalate have been trialled, with limited success, as well as mixed acid electrolytes (HCL/H_2SO_4, Polyacrylic/methanesulfonic acid), both of which were shown to improve solubility, thermal stability.[10,20] Sulfatechloride mixed electrolytes were particularly good, shown to allow 2.5M vanadium, resulting in a near 70% energy capacity increase over the pure sulfate system and giving an operational range of 5 to +50 C.[16]

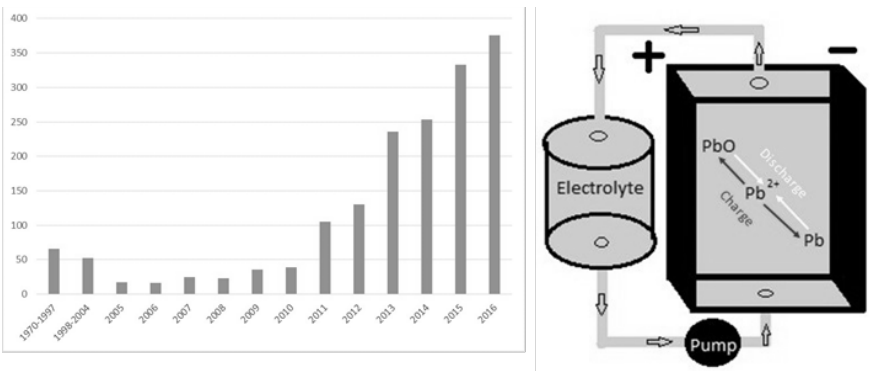


Figure 2:
A) Number of published papers on redox flow batteries by year to 2016 (as derived from ISI web of science)

B) Schematic of a lead flow battery

The current cost of Vanadium Redox Flow cells is estimated to be greater than $180/kWh even at grid level scale production[11] making it economically infeasible for large scale grid applications in its current state.[8] Much of this is due to the cost of vanadium (costing ≈ $90/kWh at its current price).[22] This makes the challenge

largely about extending the life of the cell to spread the cost over time. The development of selective (although more expensive) ion-membranes to prevent crossover between the two cells has been shown to greatly increase efficiency over time and thus battery lifespan. Both Sulfonated poly(ether ether ketone) (SPEEK) membranes embedded with graphene and Solgel-derived Nafion/SiO_2 hybrid membranes have been shown to increase coulombic and energy efficiencies while decreasing crossover of vanadium ions.[23–25] Combining the two, making a membrane of ultra-thin Nafion film on the PES/SPEEK porous membranes was found to give very high ion selectivity and energy efficiency (86.5% at 80 mA/cm^2).[26] However these membranes are very expensive, and their added cost currently outweighs the benefit of the longer lifespan.

It is also possible to reduce the cost of the RFB by reducing material cost. Most of an RFB is made of largely "off the shelf" components, meaning that material cost reduction can only be done by changing cell stack and active components.[8,12] Many paths are being explored, however a promising idea is to replace the catholyte (V^{4+}/V^{5+}) with an Fe^{2+}/Fe^{3+} catholyte,[27] significantly reducing the amount of Vanadium in the cell and thus reducing the cost. The cell has been shown to be stable from 0−50°C with a voltage range of 0.5-1.35V. Further research is ongoing.[10] There is also a lot of research looking into cost reduction by performance increase using the same or similarly priced materials (reducing kg/kWh).[12]

Significant performance improvement has been seen by the use of advanced electrode materials and design. The use of nanofibres is seen to double the electrode surface area to 350cm^2/cm^3, allowing greater power,[28] and the use of microwave treated graphite felt (at 400°C) has been shown to improve electrocatalytic activity and efficiency.[29] This is part of a general trend, with nanostructured electrodes having been shown to greatly increase performance (especially efficiency, speed of discharge) in multiple battery systems.[30–32] A recent improved cell design (applicable to all flow batteries) has also allowed peak power intensity to be increased to 500mW/cm^2,[33] allowing greater power density to be achieved.

Energy density can also be improved using V Br_3 and V Br_4 (1:1) in HBr as the electrolyte (both half cells), which allows vanadium concentration to be increased to 4 M and allows transfer of 1.5 electrons per V atom at the anode as well as increasing capacity at the cathode due to bromide to bromine oxidation.[34] This substantially increases energy density, but results in a relatively low coulombic efficiency of 88% and the production of toxic bromine in the catholyte.[10]

Recently, significant progress has been made with the demonstration of systems of MWh capacity,[5,35] however it is still clear that Vanadium RFB technology is not at a level of performance/cost effectiveness for large scale deployment.[10,15] Despite the advances outlined, the high cost of Vanadium and the restrictive temperature limits remain an issue. The chemistry of the Vanadium RFB is also not yet fully understood, as there is not yet a consensus on the kinetics of the reaction, or why many of the results above improve different aspects of performance.[12,36]

Alternative Redox Flow Batteries

Vanadium cost issue has meant many researchers are looking at other active materials for use in Flow batteries. Research has been undertaken with Polysulfide-Bromine, Iron-Chromium and Metal-Bromine redox couples among others.[10] Polysulfide-Bromine was considered promising due to the low cost of electrolyte, however cross contamination (resulting in low efficiency), slow ion transport (sodium) and high potential for environmental pollution has meant research has largely diverted elsewhere.[13] Iron-chromium (Fe^{3+}/Fe^{2+}––Cr^{3+}/Cr^{2+} redox couples) has also long been considered, and a 1MWh demonstration battery was recently installed,[10,37] with energy efficiency of 70-80% . The main advantage of such a system is that Iron and chromium electrolytes are significantly cheaper and environmentally benign. However lower voltage output (1.18V), lower current density and a high capacity decay rate per cycle (1.2%) are all drawbacks.[38] There is also the additional requirement for a heat exchanger with an iron-chromium couple. Iron-chromium systems are currently initially cheaper than Vanadium systems per kWh,[38] however their high decay rate is of concern and significant development to increase current density and battery life is required if they are to become cost-effective. In particular, a more selective membrane between cells is needed.[10,38] Metal-Bromine batteries mean replacing the anolyte with a metal anode. In this area, the zinc-bromine RFB is the most developed,[10] with an average cell voltage of 1.67V. The use of a $Br/ClBr_2$ redox couple with a Zn/Zn^{2+} anode was found to increase energy efficiency to 81%, with a Coulombic efficiency of 96%.[39] The technology is particularly cost effective in remote telecom applications[10] and a number of companies continue to pursue its development.

Another area of interest is all metal flow batteries, such as the all lead battery, which uses the redox couple described by the half cell reactions in 3,4, charging in the forward direction. This system does not require separation of anolyte and catholyte,[40] as is seen in Figure 2b, giving it a significant advantage over a Vanadium RFB in not needing an exchange membrane. The relative cheapness of the active materials and the simple design are also promising. However the cycle life of Lead RFBs (and in fact Lead-acid batteries) has traditionally been their biggest drawback.[13] This problem has now been solved somewhat by optimizing PbO_2 nanoparticle depositions on the positive electrode, which increases the battery cycle life to over 2000 cycles at 79% energy efficiency.[40] This is an important breakthrough and should generate further interest in the area, however the technology is currently far from mature.

Negative electrode: $Pb^{2+} + 2e^- \leftrightharpoons Pb$ (−0.13V) (3)

Positive electrode: $Pb^{2+} + 2H_2O \leftrightharpoons PbO_2 + 4H^+ + 2e^-$ (+1.46V) (4)

There have also been some attempts to develop a flow battery with a non-aqueous electrolyte, (redox reaction does not involve water, meaning the ion exchanged between catholyte and anolyte is usually not a proton, making movement

slower). Some of these devices have shown promising cell potentials (>2V), but low solubility of complexes, low ionic conductivity and low efficiencies have hampered development.[41] Despite the general trend, some nonaqueous RFBs appear promising. One such battery uses a $[Co(bpy)^3]^{+/2+}$ anolyte and a $[Fe(bpy)^3]^{2+/3+}$ catholyte in 0.5M solution of NEt_4BF_4 in propylene carbonate, carbon-coated Ni-FeCrAl and Cu metal foam electrodes, giving a working voltage >2.1 V and 85% energy efficiency without diminishment over 300 cycles.[10] More research into the area may yet produce results, however the higher solvent costs must be offset.[12]

It is also worth noting the emergence of Organic flow batteries, making use of organic electroactive materials, in both aqueous and non-aqueous solvents. Quinone/hydroquinone couples are being considered as materials.[42] These have the advantage of low cost, good solubility in various electrolytes, and an opportunity to tune redox properties, crossover transport by introducing molecule substituents.[10] Use of non-aqueous solvents allow higher energy densities to be achieved, but has shown stability (lifetime) issues. Recent physical-chemical modelling has however shown promise in finding highly efficient electrolytes with much lower capacity loss (undetectable over 200 cycles)[21], though the high expense of such electrolytes remains an issue.

This is found to be in common with lithium flow batteries, which can be made quite energy dense, but remain quite costly. Lithium flow batteries however, are still receiving a significant amount of research attention.[11] Researchers hope to use our relatively mature chemical understanding of lithium ion chemistry alongside the greatest advantage of flow batteries: energy and power independence, to produce a battery of grid scale potential.[11] A semi flow battery type dominates research, which differs from "full flow" batteries by using a redox active liquid (slurry) in place the usual system of dissolving active materials in electrolytes.[11] A disadvantage to such a system is the high viscosity of the slurries leading to high energy losses in the pump.[10] This is somewhat balanced out by the extremely high coulombic energy efficiency that can be achieved (99%) with energy density of 63Wh/kg (nearly 3 times Vanadium RFBs).[43] High capacity fading (40% in 100 cycles) however, suggests that the technology is also still far from ready to challenge the more mature Lithium Ion and Vanadium RFB batteries.

Conclusions

Flow Batteries are one of the most promising solutions to the problem of large scale energy storage due to their simple design, low cost and the uncoupling of power and energy allowing tailored applications. Research in the area has accelerated rapidly in the last few years as researchers have recognised its potential. Vanadium flow batteries are currently the most mature and market ready technology, with a significant amount of effort currently dedicated to their improvement toward a target of $120/kWh. While this effort is ongoing, there is also significant promise

being shown in alternative flow battery technologies such as lead, lithium and zinc-bromine. However no flow battery technology has yet shown enough promise to make large scale energy storage economical. With the use of new materials and modern nanoscale techniques, technologies such as Vanadium and lead RFBs are getting closer though. It also remains extremely possible that the most promising materials have yet to be found, with vast areas such as organic and non-aqueous flow batteries only in their infancy. As these fields develop, it should be ensured that focus is kept on pathways to lower cost, as this is ultimately what will allow wide-scale adoption for grid energy storage. Long battery lifetimes are also important, in order to minimize maintenance/part replacement costs, and methods of determining lifetime in a lab must be improved for flow technology, requiring a much better understanding of the mechanisms which currently cause degradation. The flow battery is yet far from fully developed or understood. This provides enormous opportunity for research, which if successful may well lead to a large scale energy storage solution, full integration of renewables to the grid and a secure energy future.

Acknowledgements

I would like to acknowledge the help of Stephen Byrne for his help in reviewing the paper.

References

1. IEA, *Key World Energy Statistics* 2016. Available at https://www.iea.org/publications/freepublications/publication/key-world-energy-statistics.html. (a): p6, 58 (conversion) (b):p46 NPS, 58 (conversion)

2. D. Larcher, J.M Tarascon. *Nature Chemistry* 2015, 7, 19–29.

3. United Nations Paris Agreement, (C.O.P. 21). 12 December 2015.

4. Malte MeinHausen, S. J Smith, K. Calvin, J.S Daniel, M.L.T Kainuma, J-F. Lamarque, K. Matsumoto, S.A Montzka, S.C.B. Raper, K. Riahi, A Thomson, G.J.M.Velders, D.P.P van Vuuren *Climatic Change* 2011,

5. Xing Luo, Jihong Wang, Mark Dooner, Jonathan Clarke, *Applied Energy* 2015, 137, 511–536.

6. Haisheng Chen, Thang Ngoc Cong, Wei Yang, Chunqing Tan, Yongliang Li, Yulong Ding *Progress in Natural Science* 2009, 19, 291–312.

7. T Gasser, C Guivarch, K. Tachiiri, C.D Jones, P Ciais *Nature Communications* 2015,

8. Robert M. Darling, Kevin G. Gallagher, Jeffrey A. Kowalski, Seungbum Ha and Fikile R. Brushett *Energy Environmental Science* 2014, 7, 3459.

9. Bruce Dunn, Haresh Kamath, Jean-Marie Tarascon *Science* 18 Nov 2011, 334, 928–935.

10. Grigorii. L. Soloveichik, *American Chemical Society* 2015, 115, 11533–11558.

11. Yu Zhao, Yu Ding, Yutao Li, Lele Peng, Hye Ryung Byon, John B. Goodenough and Guihua Yu *Chem Soc Review* 2015, 44, 7968.

12. Mike L. Perry, Adam. W. Zeber. Journal

of the *Electrochemical Society* 2016, 5064–5067.

13. edited by Chris Menictas, M. S.-K.; Lim, T. M. *Advances in Batteries for Medium and Large Scale Energy Storage*; Woodhead Publishing Series in Energy, 2014; multiple authors.

14. Hyun Deog Yoo, Elena Markevich, Gregory Salitra, Daniel Sharon, Doron Aurbach, *Materials Today* 2014, 17, 110–121.

15. Wei Wang, Qingtao Luo, Bin Li, Xiaoliang Wei, Liyu Li, Zhenguo Yang *Advanced Functional Materials* 2012, 23, 970–986.

16. Li, Liyu; Kim, Soowhan; Wang, Wei; Vijayakumar, M.; Nie, Zimin; Chen, Baowei; Zhang, Jianlu; Xia, Guanguang; Hu, Jianzhi; Graff, Gordon; Liu, Jun; Yang, Zhenguo *Advanced Energy Materials* 2011, 1, 394–400.

17. Thaller, L.H *J. Electrochemical Soc.* 1977, 124, C278–279.

18. E. Sum, M. Skyllas-Kazacos. *J. Power Sources* 1985, 15, 179.

19. M. Skyllas-Kazacos, R. A.-M., M.Rychcik *J. Electrochemical Society* 1986,

20. Zhang, Jian-Lu; Li, Li-Yu; Nie, Zi-Min; Chen, Bao-Wei; Vijayakumar, M.; Kim, Soow-Han; Wang, Wei; Schwenzer, Birgit; Liu, Jun; Yang, Zhenguo, J. Applied Electrochemistry 2011, 41, 1215–1221.

21. Sevov, Christo. S; Hickey, David P.: Cook, Monique E.; Robinson Sophia G. Barnett, Shoshanna; Minteer, Shelley D.; Sigman, Matthew S; Sanford, Melanie S. *J American Chemical Society* 2017

22. Gareth Kear, Frank C. Walsh., Akeel A. Shah J. *Energy Research* 2012, 36, 1105–1120.

23. Xi, Jingyu; Wu, Zenghua; Qiu, Xinping; Chen, Liquan *J. Power Sources* 2007, 166, 531–536.

24. Xi, Xiaoli; Ding, Cong; Zhang, Hongzhang; Li, Xianfeng; Cheng, Yuanhui; Zhang, Huamin. *Power Sources* 2015, 274, 1126–1134.

25. Dai, Wenjing; Yu, Lihong; Li, Zhaohua; Yan, Jia; Liu, Le; Xi, Jingyu; Qiu, Xinping *Electrochimica Alta* 2014, 132, 200–207.

26. Li, Yun; Li, Xianfeng; Cao, Jingyu; Xu, Wanxing; Zhang, Huamin *Chemical Communications* 2014, 50, 4596–4599.

27. Wang, Wei; Kim, Soowhan; Chen, Baowei; Nie, Zimin; Zhang, Jianlu; Xia, Guan-Guang; Li, Liyu; Yang, Zhenguo; Kim *Energy and Environmental Science* 2011, 4, 4068–4073.

28. Flox, Cristina; Fabrega, Cristian; Andreu, Teresa; Morata, Alex; Skoumal, Marcel; Rubio-Garcia, Javier; Morante, Juan Ramon *RSC Advances* 2013, 3, 12056–12059.

29. Wu, Xiaoxin; Xu, Hongfeng; Xu, Pengcheng; Shen, Yang; Lu, Lu; Shi, Jicheng; Fu, Jie; Zhao, Hong *J. Power Sources* 2014, 263, 104–109.

30. Faxing Wang, Xiongwei Wu, Chunyang Li, Yusong Zhu, Lijun Fu, Yuping Wu, Xiang Liu *Energy and Environmental Science* 2016, 9, 3570.

31. Yu Zhao,ab Yu Ding, Yutao Li, Lele Peng, Hye Ryung Byon, John B. Goodenough and Guihua Yu. *Nano Letters* 2014,

32. Y. Munaiah, S. Suresh, S. Dheenadayalan, Vijayamohanan K. Pillai, and P. Ragupathy *J. Phys Chemistry* 2014, 118, 14795–14804.

33. Perry, M. L.; Darling, R. M.; Zaffou, *R. ECS Transactions* 2013, 53, 7–16.

34. Kazacos, M. K. N., M. Skyllas-Kazacos *U.S Patent* 7976974 2011,

35. M. Skyllas-Kazacos, M. H. Chakrabarti, S. A. Hajimolana, F. S. Mjalli, M. Saleem *J. Electrochemical Society* 2011, 158, R55–R79.

36. N. Pour, D. G. Kwabi, T. Carney, R. M. Darling, M. L. Perry, and Y. Shao-Horn *J. Physical Chemistry* 2015, 119, 5311.

37. EnerVault, www.enervault.com,
38. Y.K Zeng, T.S. Zhao, L. An, X.I. Zhou, L. Wei *J. Power Sources* 2015, 300, 438–443.
39 Zhang, Liqun; Lai, Qinzhi; Zhang, Jianlu; Zhang, Huamin *ChemSusChem* 2012, 5, 867–869.

40. Michael G. Verde, Kyler J. Carroll, Ziying Wang, Aaron Sathrum, Ying Shirley Meng *Energy and Environmental Science* 2013, 6, 1573.

41. Robert M. Darling, J. A. K. S. H., Kevin G. Gallagher; Brushetta, F. R. *Energy and Environmental Science* 2014, 7, 3459.

42. Yang, Bo; Hoober-Burkhardt, Lena; Wang, Fang; Surya Prakash, G. K.; Narayanan, S. R. J. *Electrochemical Society* 2014, 161, A1371–A1380.

43. Li, Zheng; Smith, Kyle C.; Dong, Yajie; Baram, Nir; Fan, Frank Y.; Xie, Jing; Limthongkul, Pimpa; Carter, W. Craig; Chiang, Yet-Ming Smith *Physical Chemistry Chemical Physics* 2013, 15, 15833.

TS
SR

Physics

TS
SR

Letter from the Editor

It is no secret that the vast majority of serious scientific breakthroughs and discoveries emerge from the interdisciplinary shadows; a place where traditionally defined boundaries blur and mix affording scientists the opportunity to gaze deeper into the clockwork of our world than previously thought possible. It is this very insight that has the potential to seismically affect how our society orientates itself in an increasingly complex and obscured future.

This sentiment is echoed by the European Commission, which recently outlined six Key Enabling Technologies(KET's) it feels are tantamount to the creation of a solid and economically sustainable industrial base from which European society can develop and prosper. Of these technologies, half are reliant upon the kind of nanoscale technology that Trinity has become known for. Trinity has long established itself as a hub of nanoscale research and innovation, with the CRANN research institute continuing its charge as a world leader in KET's. Trinity's physics graduates are, therefore, uniquely well placed to take advantage of this interdisciplinary renaissance which will see our planet and its constituent societies prosper long into the future.

In this, the third year of the journal's operation, we have received more submissions than ever and each treated their chosen topic with a delicacy and intuition befitting the most skilled of scientific writers. The quality of this journal is a testament not only to the editorial team and participating students but to the wider student base as a whole. The fact that there is such a hunger for the organisation and dissemination of scientific knowledge is inspiring and presages hope for the future of rationalism.

In closing I would like to thank Prof. I. Shvets, Mr. A. Kelly, Mr. S. Roychoudhury, Mr J.Boland and Mr. S. O'Brien for their assistance and advice in ensuring the scientific rigour of this publication. Their input is invaluable and forms the backbone of the peer review process without which scientific progress would be impossible. It is my hope that you will enjoy and learn from the papers that follow, each of which sits squarely on the theme discussed in this editorial, that of progress through collaboration.

Jeffrey Lyons
Physics Editor
Trinity Student Scientific Review 2017

TS
SR

Perovskite Tandem Solar Cells: If You Can't Beat Them, Adjoin Them

Peter Foley
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

Although existing photovoltaic (PV) solar cell technologies are available at an ever decreasing price to the consumer, stagnating efficiencies necessitate a superior alternative. Perovskite cells are the fastest improving solar technology yet realised, reaching efficiencies silicon cells took 30 years to achieve, in just 3 years. Perovskite tandem solar cells offer a promising avenue for enhanced performance with increased cost effectiveness. This review focuses on the emerging technology of perovskite solar cells, assessing the current performance and future potential of three promising cell types: The single junction perovskite solar cell, the perovskite-silicon tandem solar cell and the perovskite-perovskite tandem solar cell. The single junction perovskite cell has already matched current PV efficiencies yet a viable mass production method has to be achieved. Perovskite-silicon tandem solar cells can be formed by adding perovskite to pre-installed solar modules already on rooftops. Finally, the perovskite-perovskite tandem is the most promising but furthest away from development, with the capabilities of band gap tuning leading to max efficiencies above 40% at a forecasted cost of less than half the current industry standard.

Introduction

The first evidence based global warming warning dates back to 1896 when Svante Arrhenius made the connection between CO_2 emissions and climate change.[1] Concerted efforts have been made to harness the sun's bountiful energy since the 1950's. Bell laboratories were the first to offer a vision of a society powered entirely by solar radiation when they debuted a cell that was 6% efficient in April 25, 1954.[2] In 6 decades we have witnessed efficiencies rise to 46% for some cells.

This 46% efficiency cell is a wafer bonded 4 junction Gallium based solar cell though

its high performance is outweighed by high cost.[3] It is necessary for a solar cell to fulfill certain criteria, namely, high efficiency coupled with low cost and a mass producible design. The silicon photovoltaic (PV) cell is the current frontrunner in terms of the aforementioned criteria, alas, silicon cell performance has plateaued without any inspiring increase in efficiency in many years - 25.0% to 25.6% in the previous 15 years asymptotically approaching the 29.4% Auger-recombination-constrained Shockley-Queisser limit, which is the theoretical maximum efficiency a single junction p-m solar cell can reach.[4,5] Perovskite solar cell technology has displayed rapid improvements and is currently the fastest progressing solar technology ever developed, achieving efficiencies matching silicon based cells in only 10% of the development time.[5] It was initially thought for single junction perovskite cells to best and overthrow current silicon cells, this is a long term goal in solar technology. However a more immediate alternative is needed as silicon PV cells over time have become too cheap to compete with directly. At the PSCO 2016 - the 2nd international conference on perovskite solar cells and optoelectronics held in Genoa – an alternative was found. The aim decided upon was to compliment established silicon technology with perovskite by forming a tandem cell design with the addition of a perovskite layer to a silicon solar cell to boost power outputs significantly.

A tandem cell consists of two or more sub cells, each sub cell containing an absorbing layer with a distinct band gap. By combining these sub cells, a tandem cell can produce a higher voltage than its single-junction counter part. There are two types of tandem: 2 terminal/monolithic or 4 terminal/mechanically stacked. In a 2 terminal tandem the cells are connected in series whereas in a 4 terminal tandem the sub cells are independent of one another. As the 2 terminal cells are connected in series the voltages add but the current is limited by the lowest sub cell. A 4 terminal tandem by definition has four electrodes, 3 of which must be transparent. Optical losses will be present however as full transparency is not achieved. A 4 terminal tandem is not limited by its lowest sub-cell and can also be retrofitted to pre-existing solar cells with relative ease. Through tuning the bandgaps of each absorbing material one can overcome the Shockley-Queisser limit as the bottom cell can absorb low energy photons passed through the top cell leading to higher voltages and therefore higher power conversions.

To understand bandgap tuning we must look at the composition of perovskite. Metal halide perovskites are of the form ABX_3: where A is typically methylammonium (MA), or formamidinium (FA) or Cs; B is Pb or Sn; and X is I or Br. The bandgap of perovskite can range from 1.2eV – 2.3 eV.[6–8] In-keeping with the general formula, by replacing all X from Iodine to bromine in a perovskite of $MAPbI_3$ to make $MAPbBr_3$ the bandgap can be shifted from 1.6eV to 2.3eV. The A site cation affects the band gap based primarily on its size. The ideal perovskite metal halide framework has an octahedral structure and an 180° bond angle between Pb-I-Pb. The changing of the A site cation causes adjacent octahedra to tilt with respect to one another, this reduces the Pb-I-Pb bond angle. The bonds are a hybrid of halide p orbitals and metal s orbitals, this interaction is directionally dependent and any deviation

from 180° results in worse overlap of atomic orbitals. The top of the valence band consists of halide p and metal s antibonding orbitals. The decreased overlap due to the tilting reduces the bonding-antibonding splitting between the metal and the halide, thus lowering the energy of the valence band maximum and increasing the bandgap. For B the use of Sn instead of Pb will lower the bandgap and varying the compositions of any of these in ratio can reach intermediate values all with varying stabilities.

Single-Junction and Perovskite Tandem cells

Single-Junction Perovskite Solar Cells

Exhibiting rapidly increasing power conversion efficiencies (PCEs) and low production costs, singlejunction perovskite devices have reached a certified 22% PCE.[5] Whilst initially developed to outperform existing technology it has come to pass that the first commercial iterations of perovskite PVs will likely be as an "add-on" to silicon (Si) PVs to form tandem cells. A perovskite solar cell SEM seen in the figure below displays each intricate layer of the cell.

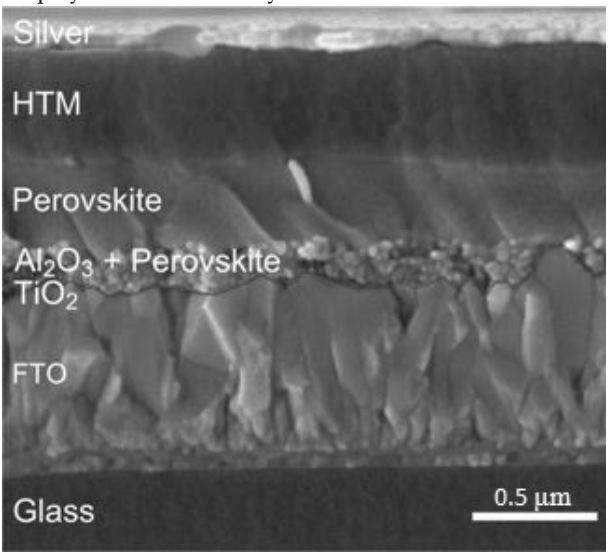


Figure 1: Each intricate layer of a perovskite cell as viewed under a SEM. From the bottom up: Fluorine doped tin oxide (FTO), Titanium oxide, aluminium oxide with perovskite, perovskite, the holetransporting material (HTM) consisting of poly(triaryl amine) (PTAA) and finally the opaque back contact - silver.
Figure adapted from Ball et al.[9]

A solar cell's capabilities are defined by the following characteristics: Short circuit current (JSC), open circuit voltage (VOC), Fill Factor (FF) and power conversion efficiency (PCE %). A characteristic J-V curve of current density versus voltage is the acid test in scrutinising the efficiency of a solar cell. For the most efficient perovskite solar cell in its class the following specifications have been produced in a lab, see also fig. 3 for the J-V and PCE curves for the champion cell fabricated by Saliba *et al*.[10]

	J_{sc} (mA cm^{-2})	V_{oc} (mV)	FF	PCE (%)
Saliba *et al.*[10]	22.69 ± 0.75	1132 ± 25	0.748 ± 0.018	21.1
Saliba *et al.*[11]	22.8	1180	0.81	21.6
Gratia *et al.*[12]	23.2	1110	0.79	20
Jeon *et al.*[13]	22.5	1110	0.73	18.4

Table 1: Figures of merit for the four leading single-junction perovskite solar cells.

The precise architecture and fabrication methods of each cell are the cause of the differing efficiencies. The lowest efficiency of the table 18.4% Jeon *et al*. and 20 % Gratia *et al*. both use a perovskite composition of $(FAPbI_3)_{0.85}(MAPbBr_3)_{0.15}$ with identical solar cell architecture, an electron contact of mesoporous TiO_2 and a hole contact of PTAA but the processing conditions have been optimised for the latter. This highlights the importance of optimal device fabrication parameters. For the cell of 21.6% efficiency, a composition with 4 cations[11] was used whilst a composition of 3 cations[10] used for the cell of 21.1% efficiency. Both of these cells use a Spiro-OMeTAD hole contact layer. This increased efficiency in the 4 cation cell is due to additional Rb^+ which alters the film formation kinetics which they suggest results in the formation of a more photo-active perovskite phase.

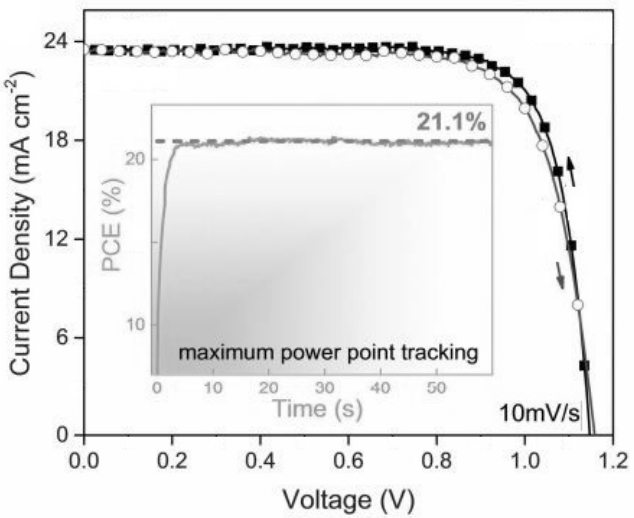


Figure 2: J-V curves of Saliba *et al.* champion single-junction perovskite cell. The cell was illuminated under light intensity equivalent to that of one sun. The J-V curves were measured by sweeping the bias from positive to zero indicated by the arrow and again from zero to positive (only a minor amount of hysteresis on the return measurement). The y-intercept represents a J_{sc} close to 24 mA $cm^{-2}$and x-intercept a V_{oc} of 1.13 V. A PCE of 21.1% is seen within the graph, this is obtained by finding the max product of voltage and current and measuring performance over an extended time period.
Figure adapted from Saliba et al.[10]

Perovskite-Silicon Tandems

Perovskite solar cells have the desirable trait of a tunable bandgap. This characteristic is massively advantageous and is crucial in cementing its ability to function as a member of a tandem cell. With a bandgap tunable within the range of 1.2 eV - 2.3 eV,[6–8] it would be desirable to mount a top perovskite cell of bandgap 1.7 to 1.9 eV on a Silicon cell of band gap 1.1 eV.[14]

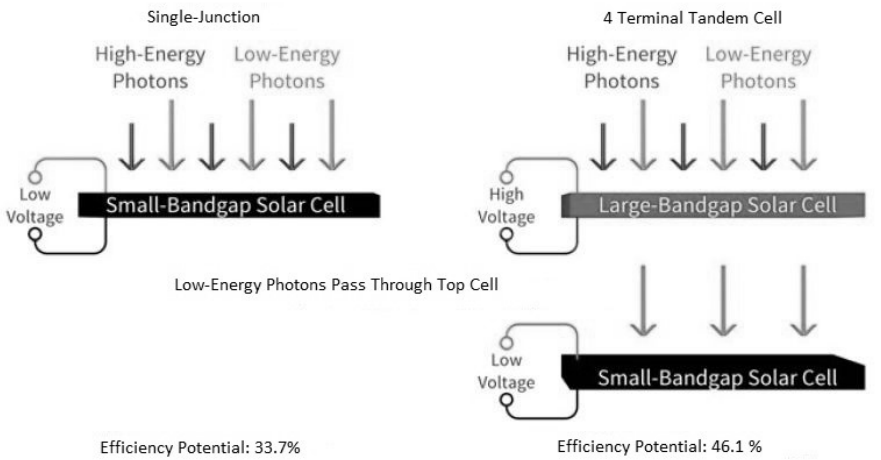


Figure 3: A comparison of theoretical max efficiencies for a single-junction and tandem solar cell is seen in this figure. Although the single-junction cell is capable of capturing both high-energy and low-energy photons due to its low bandgap it does so at a much lower voltage than the tuned large band gap tandem solar cell. Increased voltage means higher energy output with a theoretical increase of 12.4%.
Figure was adapted from Bailie et al.[15]

A number of groups are working on the promising Perovskite-Silicon tandem cells with much success. Results are displayed in Table 2 below:

	J_{sc} (mA cm^{-2})	V_{oc} (mV)	FF	PCE (%)
Albrecht *et al.*[16]	14	1.78	0.79	19.9
Werner *et al.*[17]	16.1	1.70	0.70	19.5
Mailoa *et al.*[18]	11.5	1.58	0.75	13.7
Bush *et al.*[19]	18.09	1.651	0.79	23.6 ± 0.6

Table 2: Figures of merit for leading perovskite-silicon tandem solar cells.

The secret to the efficiency of the cell fabricated by Albrecht *et al.* is the presence of amorphous silicon layers either side of the crystalline silicon. The amorphous layers passivate surface defects known to occur in crystalline silicon. This heterojunction architecture is not present in the cell fabricated by Mailoa *et al.*, a more standard p-n junction is present in its place. Two subtle differences are observed between the first two cells of the table. For the Albrect *et al.* cell recombination layers of Indium tin oxide (ITO) are used while in the Werner *et al.* cell an ITO with Indium zinc oxide (IZO) is utilised. Also the electron selective contact differs in both, with a SnO_2 layer for the top cell and a fullerene (C60) layer for the second cell. The perovskite used by Albrecht *et al.* is $(FAPbI_3)_{0.85}(MAPbBr_3)_{0.15}$ with a bandgap of 1.56 eV while Werner *et al.* and Mailoa opted for $MAPbI_3$ with a bandgap of 1.57 eV all of which are far below the ideal bandgap of 1.75 eV leaving large room for improvement upon finer tuning, particularly to the J_{sc}. The current world record is held by the cell produced by Bush *et al.* however the structural details of the cell have yet to be published, only specifications are known.

Perovskite-Perovskite Tandems

This specific tandem has the potential to significantly reduce costs and increase practicalities of the cell fabrication process. While the silicon layer of a perovskite-silicon tandem would require high temperature, high purity solid state processing, both sub cells of the perovskite-perovskite tandem could be fabricated at lower cost.

Perovskite-Perovskite Tandems have been made with an upper cell of bandgap 1.8 eV and lower cell of bandgap 1.2 eV with impressive results.[8]

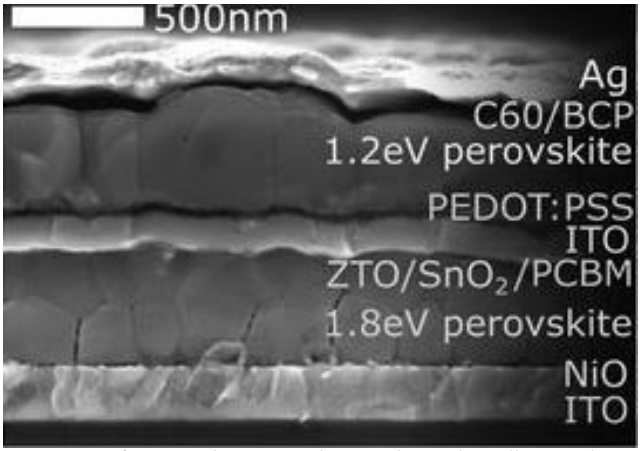


Figure 4: SEM image of a perovskite-perovskite tandem solar cell. Note the tuned bandgaps of the upper and lower perovskite layers. From the bottom up we see indium tin oxide (ITO), Nickel Oxide, a 1.8eV perovskite layer, a combined Zinc tin oxide (ZTO), Tin Oxide and phenyl-C_{61}-butyric acid methyl ester (PCBM) layer, ITO, poly(3,4-ethylenedioxythiophene) (PEDOT:PSS), a 1.2eV perovskite layer, C60 with bathocuproine (BCP) and finally the opaque silver back contact.
Figure adapted from Eperon et al.[8]

The cell produced by Jiang *et al.* was an early perovskite-perovskite tandem and therefore did not amaze in terms of efficiency. It consists of a tandem of two identical perovskites $MAPbI_3$ so bandgap optimisation is not present. The top contact for holes used was a Spiro-OMeTAD and a bottom contact of PCBM:PEI for electron contact. This cell is mainly a proof of concept and that a tandem perovskite-perovskite solar cell could exist. The Eperon *et al*. fabrication uses a 1.2 eV bottom cell of $FA_{0.75}Cs_{0.25}Sn_{0.5}Pb_{0.5I3}$ and a top cell of $FA_{0.83}Cs_{0.17}Pb(I_{0.5}Br_{0.5})_3$ with a bandgap of 1.8 eV. For the Forgacs *et al.* cell the efficiency is impressively high for a tandem cell who's bandgaps are not optimised. The tandem consists of a $MAPbI_3$ bottom cell and $FA_{0.85}Cs_{0.15}Pb(I_{0.3}Br_{0.7})_3$.

	J_{sc} (mA cm^{-2})	V_{oc} (mV)	FF	PCE (%)
Jiang *et al.* (2T)[20]	6.61	1890	0.56	7.0
Forgacs *et al.*(2T)[21]	9.83	2294	0.803	18.1
Eperon *et al.*(2T)[8]	14.5	1660	0.70	17.0

Table 3: Figures of merit for several recently reported perovskite-perovskite two terminal (2T) tandems.

Conclusions

Although the cost of solar energy has declined over recent years there has been little growth in efficiency of industry standard single junction solar cell. Tandem solar cells have been available for some time but their cost has been prohibitive to their widespread adoption. Perovskite tandem solar cells represent a viable option for high efficiency low cost solar energy. Perovskite-Silicon tandems have already seen high efficiencies as silicon is a long established technology and represents a medium term option in solar energy. Perovskite-Perovskite tandems, however lagging in efficiency as it stands, look to realise the long term goal of an easily produced, cost effective, high efficiency solar cell for widespread industrial and commercial use. With theoretical highs of 46% efficiency for perovskite-perovskite tandems, the utility of conventional PV cells may be severely reduced. A full transposition will occur when scale dependent costs of perovskite-perovskite tandems begin to decrease.

Acknowledgements

The author would like to thank Mr. K. Frohna, Mr. J. Lyons, Mr. S. O'Brien and Prof. I. Shvets for their helpful input and appraisal.

References

1. Mark Maslin. Global Warming: A Very Short Introduction. *Oxford University Press*, 2004.

2. D. M. Chapin, C. S. Fuller, and G. L. Pearson. A new silicon pn junction photocell for converting solar radiation into electrical power. *Journal of Applied Physics*, 25(5):676–677, 1954.

3. Fraunhofer ISE. https://www.ise.fraunhofer.de/en/press-and-media/press-releases/press-releases2014/new-world-record-for-solar-cell-efficiency-at-46-percent,

4. R.M Swanson. Proceedings of 31st IEEE Photovoltaic Specialist Conference, 889, 2005.

5. NREL. Best research-cell efficiencies chart. http://www.nrel.gov/pv/assets/images/efficiency_ chart.jpg,

6. Jun Hong Noh, Sang Hyuk Im, Jin Hyuck Heo, Tarak N. Mandal, and Sang Il Seok. Chemical management for colorful, efficient, and stable inorganic–organic hybrid nanostructured solar cells. *Nano Letters*, 13(4):1764–1769, 2013. PMID: 23517331.

7. Giles E. Eperon, Samuel D. Stranks, Christopher Menelaou, Michael B. Johnston, Laura M. Herz, and Henry J. Snaith. Formamidinium lead trihalide: a broadly tunable perovskite for efficient planar heterojunction solar cells. *Energy Environ.* Sci., 7:982–988, 2014.

8. Giles E. Eperon, Tomas Leijtens, Kevin A. Bush, Rohit Prasanna, Thomas Green, Jacob Tse-Wei Wang, David P. McMeekin, George Volonakis, Rebecca L. Milot, Richard May, Axel Palmstrom, Daniel J. Slotcavage, Rebecca A. Belisle, Jay B. Patel, Elizabeth S. Parrott, Rebecca J. Sutton, Wen Ma, Farhad Moghadam, Bert Conings, Aslihan Babayigit, Hans-Gerd Boyen, Stacey Bent, Feliciano Giustino, Laura M. Herz, Michael B. Johnston, Michael D. McGehee, and Henry J. Snaith. Perovskiteperovskite

tandem photovoltaics with optimized bandgaps. *Science,* 2016.

9. James M. Ball, Michael M. Lee, Andrew Hey, and Henry J. Snaith. Low-temperature processed meso-superstructured to thin-film perovskite solar cells. *Energy Environ.* Sci., 6:1739–1743, 2013.

10. Michael Saliba, Taisuke Matsui, Ji-Youn Seo, Konrad Domanski, Juan-Pablo Correa-Baena, Mohammad Khaja Nazeeruddin, Shaik M. Zakeeruddin, Wolfgang Tress, Antonio Abate, Anders Hagfeldt, and Michael Gratzel. Cesium-containing triple cation perovskite solar cells: improved stability, reproducibility and high efficiency. *Energy Environ.* Sci., 9:1989–1997, 2016. 11Michael Saliba, Taisuke Matsui, Konrad Domanski, Ji-Youn Seo, Amita Ummadisingu, Shaik M. Zakeeruddin, Juan-Pablo Correa-Baena, Wolfgang R. Tress, Antonio Abate, Anders Hagfeldt, and Michael Gr¨atzel. Incorporation of rubidium cations into perovskite solar cells improves photovoltaic performance. *Science,* 354(6309):206–209, 2016.

12. Paul Gratia, Giulia Grancini, Jean-Nicolas Audinot, Xavier Jeanbourquin, Edoardo Mosconi, Iwan Zimmermann, David Dowsett, Yonghui Lee, Michael Gr¨atzel, Filippo De Angelis, Kevin Sivula, Tom Wirtz, and Mohammad Khaja Nazeeruddin. Intrinsic halide segregation at nanometer scale determines the high efficiency of mixed cation/mixed halide perovskite solar cells. *Journal of the American Chemical Society,* 138(49):15821–15824, 2016. PMID: 27960332.

13. Nam Joong Jeon, Jun Hong Noh, Woon Seok Yang, Young Chan Kim, Seungchan Ryu, Jangwon Seo, and Sang Il Seok. Compositional engineering of perovskite materials for high-performance solar cells. *Nature,* 517(7535):476–480, 01 2015.

14. Timothy J. Coutts, Keith A. Emery, and J. Scott Ward. Modeled performance of polycrystalline thin-film tandem solar cells. *Progress in Photovoltaics: Research and Applications,* 10(3):195–203, 2002.

15. Colin D. Bailie and Michael D. McGehee. High-efficiency tandem perovskite solar cells. *MRS Bulletin*, 40(8):681–686, 008 2015.

16. Steve Albrecht, Michael Saliba, Juan Pablo Correa Baena, Felix Lang, Lukas Kegelmann, Mathias Mews, Ludmilla Steier, Antonio Abate, Jorg Rappich, Lars Korte, Rutger Schlatmann, Mohammad Khaja Nazeeruddin, Anders Hagfeldt, Michael Gratzel, and Bernd Rech. Monolithic perovskite/silicon-heterojunction tandem solar cells processed at low temperature. *Energy Environ.* Sci., 9:81–88, 2016.

17. J´er´emie Werner, Ching-Hsun Weng, Arnaud Walter, Luc Fesquet, Johannes Peter Seif, Stefaan De Wolf, Bjoern Niesen, and Christophe Ballif. Efficient monolithic perovskite/silicon tandem solar cell with cell area 1 cm^2. *The Journal of Physical Chemistry Letters,* 7(1):161–166, 2016. PMID: 26687850.

18. Jonathan P. Mailoa, Colin D. Bailie, Eric C. Johlin, Eric T. Hoke, Austin J. Akey, William H. Nguyen, Michael D. McGehee, and Tonio Buonassisi. A 2-terminal perovskite/silicon multijunction solar cell enabled by a silicon tunnel junction. *Applied Physics Letters,* 106(12):121105, 2015.

19. http://web.stanford.edu/group/mcgehee/research.html,

20. Fangyuan Jiang, Tiefeng Liu, Bangwu Luo, Jinhui Tong, Fei Qin, Sixing Xiong, Zaifang Li, and Yinhua Zhou. A two-terminal perovskite/perovskite tandem solar cell. *J. Mater. Chem.* A, 4:1208–1213, 2016.

21. D´avid Forg´acs, Lid´on Gil-Escrig, Daniel P´erez-Del-Rey, Cristina Momblona, J´er´emie Werner, Bjoern Niesen, Christophe Ballif, Michele Sessolo, and Henk J. Bolink. Efficient monolithic perovskite/perovskite tandem solar cells. *Advanced Energy Materials,* pages 1602121–n/a, 2016. 1602121.

TS
SR

Metamaterials and Metasurfaces:

Recent Advances and Future Outlook

Conor Coughlan
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

Recent discoveries in metamaterials and metasurfaces have singnificantly altered the way in which optical process in matter may be controlled. Light can now be manipulated in ways that differ from its propagation in natural materials. Advances such as these, coupled with new nanoscale fabrication techniques have given birth to a range of innovations and applications. Such progress is detailed in both metamaterials and metasurfaces alike, and the potential impact of such developments for imaging, holographics and thin films is discussed. Metasurfaces may prove valuable for future scientific research, through their use in tuned reflectors and absorbers, in subwavelength imaging and in generation of holograms.

Introduction

In recent years, new materials have been developed that possess unusual physical properties not observed in the natural world. These have been dubbed 'metamaterials'. The word 'meta' in Greek means 'beyond' and is so justified in the characteristics of these materials. In this review, only the resonant type of metamaterials will be discussed, specifically negative index materials (NIMs).

Metamaterials were first hypothesised explicitly by Veselago in 1967,[1] but research did not follow immediately, mainly due to a lack of understanding as to how such materials could be fabricated. He concluded that materials exhibiting negative indices of refraction could be possible if they simultaneously possessed negative real values of permittivity (ε) and permeability (μ) in the medium, which describe the electrical and magnetic response of a material. These values are related to the refractive index through $n = \pm\sqrt{\varepsilon\mu}$. It was later appreciated that a more specific relation to achieve negative indices exists[2] (given $\varepsilon = \varepsilon_1 + i\varepsilon_2$ and $\mu = \mu_1 + i\mu_2$), namely:

$$\epsilon_1|\mu| + \mu_1|\epsilon| < 0 \qquad (1)$$

This was hypothesised by substituting negative values into Maxwell's equations. Such a system was produced where the wavevector k, the magnetic field H and the electric field E form a lefthanded system. In normal media, these three vector quantities form a right-handed system. To that end, such NIMs are also known as left-handed materials. The Poynting vector may be also defined as:

$$\vec{S} = \vec{E} \times \vec{H} \qquad (2)$$

This means that in this left-handed system, the Poynting vector will have the opposite direction to the wavevector $\vec{k}$ due to the right-handed nature of the vector cross-product. A direct consequence of this is that: $\vec{S} \cdot \vec{k} < 0$. Among other corollaries are that the Doppler shift exhibited in normal materials is inverted, meaning that the frequency of light emitted from an object decreases on approach instead of increasing. The phenomenon of Cherenkov radiation is also inverted for negative refractive index materials. This allows the radiation from a charged particle to be trailing from it when travelling faster than the phase velocity of light, instead of in front of it (the case for normal media).

Different interpretations also exist for NIMs. One such class is treating the material as a photonic crystal. The crystal structure scatters the incident light in such a way that the phase velocity of the wave moves in opposite directions from the group velocity.[3] These photonic crystals were determined to have positive values of permittivity and permeability also, indicating that such a property is not universal to develop negative indices of refraction.

While theoretically possible, practical difficulties in producing materials which displayed simultaneously negative values of permittivity and permeability hampered the research into NIMs. Negative permeability values were achieved at a certain frequency for the split ring resonator (SRR) system, which consisted of two incomplete metallic rings, one inside the other, separated by a small dielectric cavity.[4] (figure 1) For permittivity, an array of thin wires in three dimensions was shown to produce negative values at definite frequencies.[5] The frequency range could be altered by changing dimensions in the wire, which changed the plasma frequency of the array. Such a system combining the two was developed a few years later, which was the world's first metamaterial.[6] The unit cell constructed would have to be a small fraction of the wavelength of incident light to achieve negative refractive indices.

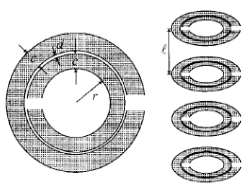


Figure 1: (A) overhead view of the SRR system (B) stacking sequency of the SRR array. Adopted from Pendry *et al.*[4]

The 'Perfect' Lens

NIMs opened the door for numerous applications, one of the main ones being in imaging. The potential for this was first realised with a 'perfect' lens.[7] Such a material could refract the light reflecting off an object so that there would be a focal point inside the material and the light rays would coincide at a point on the opposite side of the material (figure 2). However, progress towards this has been slow, owing to difficulties in producing large-scale metamaterials and adopting these materials for optical frequencies.[8]

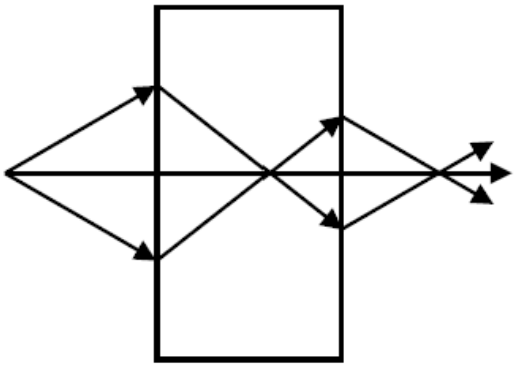


Figure 2: Negative refractive index medium focuses diverging light rays from a source. *Adopted from Pendry et al.*[7]

One of the main problems with these materials is the degree of loss of light intensity that is experienced upon transition through a layer. Therefore, very little light will be present if it goes through many millions of layers that are present on the macroscale. Eliminating these losses would allow devices such as a perfect lense to be developed. This could be achieved in two ways, either by reducing the amount of loss generated by the metamaterial, or by incorporating 'gain' materials into the nanostructure of the material. Advances in nanofabrication have allowed for the insertion of gain materials between layers such as double fishnet structures, semiconducting materials and optically pumped organic structures.[9,10] However, these technologies require much refining before being practically applicable.

Metasurfaces

If one was to scale a metamaterial down to the nanoscale, a meta-surface would result. These metasurfaces can control the phase, polarisation and therefore the amplitude of the outgoing light depending on their microstructure. Such surfaces are not subject to the vast losses in light present in metamaterials due to their two-dimensional structure. In order to be effective, the vertical dimension must only be a small fraction of the wavelength of incident light. This makes metasurfaces

with a thickness of 100 nm ideal for manipulating visible light.

Metasurfaces can be made by either optical scatterers relying on antenna dispersion or by thin films with low transmission efficiency.[11] Antenna dispersion relies on the coupling of surface plasmon polaritons with light. Plasmon polaritons are the collective electron oscillations at the interface of any two media with a real permittivity value of opposite sign (e.g. metallic-dielectric interface) coupling with light. These plasmons propagate back and forth along the surface of the media (see figure 3). The incident light interacts with the plasmons, and for a certain incident wavelength λ, antenna resonance occurs. Such resonance means that the transmitted light is in phase with the incident light. The resonance also depends on the length of the antenna or molecule used, given by:

$$L_{res} = \frac{\lambda}{2} \qquad (3)$$

If the length is smaller or larger than the resonance length, this makes the phase of the transmitted light change. Therefore, the phase of the transmitted light may be changed by altering the lengths of the antenna.

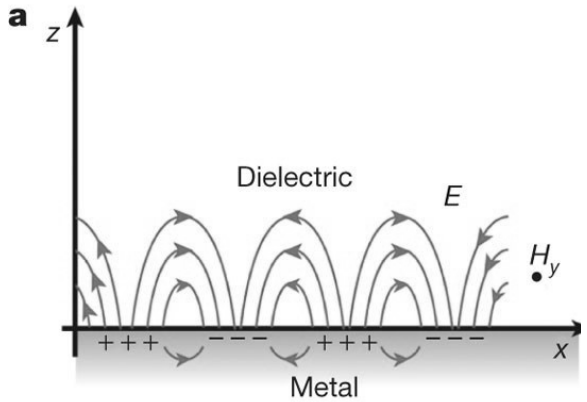


Figure 3: Description of excitation of surface plasmon polaritons at a metal-dielectric interface.
Adopted from Benson et al.[12]

The change in phase of the wave alters the direction of the propagating wave through the generalised laws of refraction and reflection.[13] These are given below respectively:

$$n_t sin(\theta_t) - n_i sin(\theta_i) = \frac{1}{k_0}\frac{d\phi}{dx} \qquad (4)$$

$$sin(\theta_r) - sin(\theta_i) = \frac{1}{n_i k_0}\frac{d\phi}{dx} \qquad (5)$$

In these equations $\theta_{i,r}$ are the angles of incidence and refraction respectively, $\frac{d\phi}{dx}$ is the phase gradient (the change in phase with respect to the depth of the surface) and k_0 is the wavevector of the incident beam in free space. This shows that light can be bent in random directions based on the phase change of the light in the medium.

Flat Lenses

By manipulation of the light source, light of subresonant-wavelength resolution may be focused down to a single point. This was first realised by generating a radial distribution of phase changes using V-shaped nanoantenna.[14] The orientation and shape of the gold nanoantenna changed with radial distance from a point, which produced different phase shifts and caused the light to focus. This was achieved for infrared wavelengths and it was shown to form a lens with a focal length *f*, the phase shift for a nanoantenna at distance (x,y) from the centre must be:

$$\phi_L(x,y) = \frac{2\pi}{\lambda}(\sqrt{(x^2+y^2)+f^2} - f) \qquad (6)$$

One drawback to this method concerns cross-polarised fields. When a wave is incident in one polarisation state, it is transmitted in its orthogonal state. Also, the transmission efficiency of the light through the surface is relatively low. One possible solution developed is a meta-transmit array.[15] A meta-transmit array utilises a stack of three thin metasurfaces made of aluminium doped zinc-oxide and a silicon dielectric. A uniform distance was assumed for each metasurface to maximise efficiency, and the phase was controlled by the ratio of zinc-oxide to the silicon (figure 4). The uniform distance allowed for the reflectance $r = 0$ and the ratio allowed $t = e^{i\varphi(x)}$ where $\phi(x)$ is the resultant phase. Simulations showed a transmission efficiency above 75% in this way.

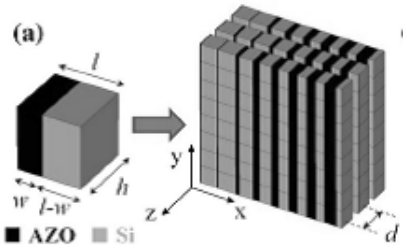


Figure 4: Meta-transmit array showing aluminium doped zinc-oxide (black) and silicon dielectric (grey).
Adopted from Monticone et al.[15]

Recently, amorphous TiO_2 nanofin deposition has lead to increased success.[16] TiO_2 was deposited on a resist in nanofins (figure 5). The phase was controlled by rotating the nanofins by a certain amount, achieving what is called a Pancharatnam-Berry phase (explained below). This was designed for specific wavelengths of 660, 532 and 405 nm, achieving a transmission efficiency of 66%, 73% and 86% respectively. These were able to resolve images containing features at subwavelength distances and achieve magnification of up to 170x.

The Pancharatnam-Berry phase alters the polarisation of light as opposed to the phase. The structure required in the metasurface is therefore different in that the scatterers are the same geometrically, but are oriented differently in space. This

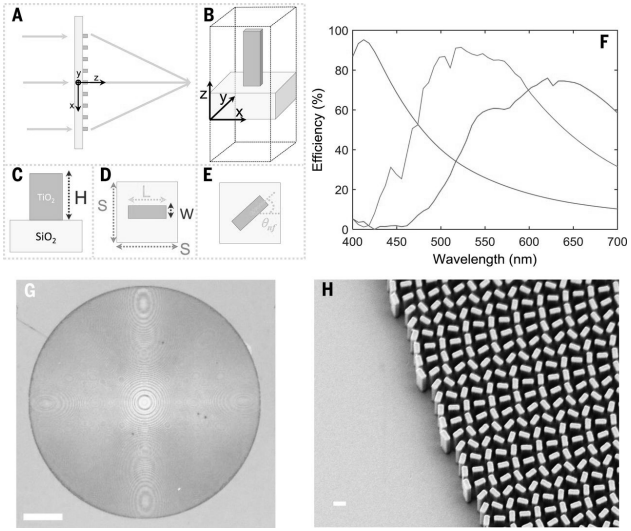


Figure 5: A-E - Schematic of lens and dimensions of the TiO_2 nanofin. F - Efficiency as a function of wavelength as per simulation of three different nanofin dimensions (In order frm left, λ = 405 nm, λ = 532 nm and λ = 660 nm). G - The metalens designed at a wavelength of 660 nm. H - SEM image of the metalens.
Adopted from Khorasaninejad et al.[16]

means that if circularly polarised light is incident, it produces a certain amount of right and left-handed polarised light, which can be controlled by altering the orientation of the scatterer. Polarisation can alter phase as shown in Jones' calculus.[11] The advantages of this method is that it can be used for broadband light as shown in figure 5 F. When just the orientation of the molecule is altered, the wavefronts are not distorted from antenna dispersion and allow for many wavelengths to be present.[11]

Holography

Achievements in the manipulation of broadband light has allowed for improvements in holographics. Incident circularly polarised light is used, and nanorods can be arranged in such a way using digital calculations to produce a 3-D image. This image may be constructed by interaction of the phases of the light transmitted.[17] This differs from conventional methods of hologram production in that multiple diffraction orders are eliminated. Advances were made in creating broadband holograms,[18] and also in adapting holograms for the three primary colours (red, blue and green).[19]

Thin film metasurfaces

As mentioned earlier, another method of metasurface production is through the use of thin films. Both the thin film and the substrate deposited have large losses. This allows very little light to travel through and reflect off the back of the substrate, meaning that the light reflected back out of the film passes through the metasurface material on top. Also, because the thin film is a metasurface, the phase of the reflected waves will be different from that of the ones that are absorbed and then reflected back out. Therefore, the waves that interact constructively or deconstructively (see figure 6) by thin film interference may be tailored by changing a number of parameters in the metasurface.This can include temperature, orientation and nature of the metasurface particles and the nature of the substrate. This in turn tunable absorbance in a thin film to be manipulated by these chosen parameters.

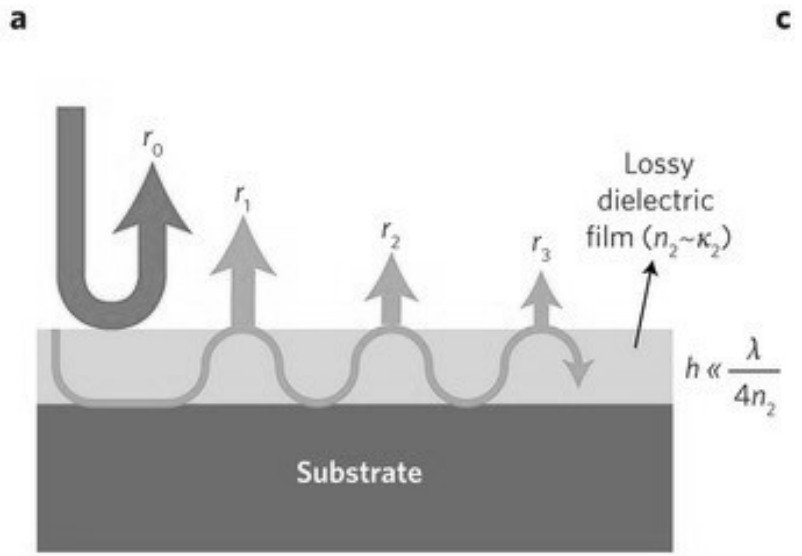


Figure 6: Thin film interference on a metasurface. The vectors r_0 –r_3 have different phases due to the metasurface. Adopted from Yu *et al.*[11]

This mechanism has been shown in a VO_2 thin film with a sapphire substrate.[20] In this case, 99.75% absorption was obtained at an incident wavelength of 11.6 μm. The absorption was altered by the temperature of the thin film, and the maximum absorption was taken at the insulator-to-metal phase transition at T = 343K.

Another proposed application is in water splitting.[21] This device uses iron oxide ($\alpha - Fe_2O_3$) as a thin film. This film is designed for minimal absorbance, and hence maximum reflectivity, the opposite of above. With this and the use of a V-shaped cell as a photon trapping device, highly efficient light harvesting could be allowed

to generate an electric current through photovoltaic methods. The current density generated was 4 mA cm^{-2}.

Discussion and Conclusions

The fabrication of a 'perfect' lens is a promising application, with potential to alter the way objects are imaged. However, this technology will take a long time to develop. Advances in nanofabrication may allow for the development of new bulk metamaterials with low losses.

Metasurfaces have undergone an incredible array of advances and innovations recently, in part due to the increasing design precision and accuracy at a nanoscale. Such technologies have proven to be useful in light harvesting and subwavelength imaging, as well as potential applications in holograph. There may also be a range of applications that have not been realised yet, particularly in thin films due to the tunable nature of absorption and reflection off of its surface.

Acknowledgements

The author would like to thank Mr. J. Lyons, Mr. S. Roychoudhury, and Prof. I. Shvets for their helpful input and appraisal.

References

1. Veselago VG. Electrodynamics of substances with simultaneously negative permittivity and permeability. *Uspekhi Fizicheskih Nauk.* 1967;92(7):517. Available from: http://ufn.ru/ ru/articles/1967/7/d/.

2. Depine RA, Lakhtakia A. A new condition to identify isotropic dielectric-magnetic materials displaying negative phase velocity. *Microwave and Optical Technology Letters.* 2004 May;41(4):315–316. Available from: http://doi.wiley.com/10.1002/mop.20127.

3. Parimi PV, Lu WT, Vodo P, Sokoloff J, Derov JS, Sridhar S. Negative Refraction and LeftHanded Electromagnetism in Microwave Photonic Crystals. *Physical Review Letters.* 2004 Mar;92(12):127401. Available from: http://link.aps.org/doi/10.1103/PhysRevLett.92. 127401.

4. Pendry JB, Holden AJ, Robbins DJ, Stewart WJ. Magnetism from conductors and enhanced nonlinear phenomena. *IEEE Transactions on Microwave Theory and Techniques.* 1999;47(11):2075–2084. Available from: http://ieeexplore.ieee.org/document/798002/.

5. Pendry JB, Holden AJ, Stewart WJ, Youngs I. Extremely Low Frequency Plasmons in Metallic Mesostructures. *Physical Review Letters.* 1996 Jun;76(25):4773–4776. Available from: http://link.aps.org/doi/10.1103/PhysRevLett.76.4773.

6. Smith DR, Padilla WJ, Vier DC, Nemat-Nasser SC, Schultz S. Composite Medium with Simultaneously Negative Permeability and Permittivity. *Physical Review Letters.* 2000 May;84(18):4184–4187. Available from: http://link.aps.org/doi/10.1103/ PhysRevLett.84.4184.

7. Pendry JB. Negative Refraction Makes a Perfect Lens. *Physical Review Letters.* 2000

Oct;85(18):3966–3969. Available from: http://link.aps.org/doi/10.1103/PhysRevLett.85.3966.

8. Soukoulis CM, Wegener M. Past achievements and future challenges in the development of three-dimensional photonic metamaterials. *Nature Photonics.* 2011 Jul;Available from: http://www.nature.com/doifinder/10.1038/nphoton.2011.154.

9. Shalaev VM. Optical negative-index metamaterials. *Nature Photonics.* 2007 Jan;1(1):41–48. Available from: http://www.nature.com/doifinder/10.1038/nphoton.2006.49.

10. Hess O, Pendry JB, Maier SA, Oulton RF, Hamm JM, Tsakmakidis KL. Active nanoplasmonic metamaterials. *Nature Materials.* 2012 Jun;11(7):573–584. Available from: http: //www.nature.com/doifinder/10.1038/nmat3356.

11. Yu N, Capasso F. Flat optics with designer metasurfaces. *Nature Materials.* 2014 Jan;13(2):139–150. Available from: http://www.nature.com/doifinder/10.1038/nmat3839.

12. Benson O. Assembly of hybrid photonic architectures from nanophotonic constituents. *Nature.* 2011 Dec;480(7376):193–199. Available from: http://www.nature.com/doifinder/ 10.1038/nature10610.

13. Yu N, Genevet P, Kats MA, Aieta F, Tetienne JP, Capasso F, *et al*. Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction. *Science.* 2011 Oct;334(6054):333–337. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1210713.

14. Aieta F, Genevet P, Kats MA, Yu N, Blanchard R, Gaburro Z, *et al*. Aberration-Free Ultrathin Flat Lenses and Axicons at Telecom Wavelengths Based on Plasmonic Metasurfaces. *Nano Letters*. 2012 Sep;12(9):4932–4936. Available from: http://pubs.acs.org/doi/abs/10. 1021/nl302516v.

15. Monticone F, Estakhri NM, Alu` A. Full Control of Nanoscale Optical Transmission with a Composite Metascreen. *Physical Review Letters.* 2013 May;110(20):203903. Available from: http://link.aps.org/doi/10.1103/PhysRevLett.110.203903.

16. Khorasaninejad M, Chen WT, Devlin RC, Oh J, Zhu AY, Capasso F. Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging. *Science.* 2016 Jun;352(6290):1190–1194. Available from: http://www.sciencemag.org/cgi/doi/10. 1126/science.aaf6644.

17. Huang L, Chen X, Mu¨hlenbernd H, Zhang H, Chen S, Bai B, *et al*. Three-dimensional optical holography using a plasmonic metasurface. *Nature Communications.* 2013 Nov;4. Available from: http://www.nature.com/doifinder/10.1038/ncomms3808.

18. Lin J, Genevet P, Kats MA, Antoniou N, Capasso F. Nanostructured Holograms for Broadband Manipulation of Vector Beams. *Nano Letters.* 2013 Sep;13(9):4269–4274. Available from: http://pubs.acs.org/doi/abs/10.1021/nl402039y.

19. Huang YW, Chen WT, Tsai WY, Wu PC, Wang CM, Sun G, *et al*. Aluminum Plasmonic Multicolor Meta-Hologram. *Nano Letters*. 2015 May;15(5):3122–3127. Available from: http: //pubs.acs.org/doi/abs/10.1021/acs.nanolett.5b00184.

20. Kats MA, Sharma D, Lin J, Genevet P, Blanchard R, Yang Z, *et al*. Ultra-thin perfect absorber employing a tunable phase change material. *Applied Physics Letters*. 2012 Nov;101(22):221101. Available from: http://aip.scitation.org/doi/10.1063/1. 4767646.

21. Dotan H, Kfir O, Sharlin E, Blank O, Gross M, Dumchin I, *et al*. Resonant light trapping in ultrathin films for water splitting. *Nature Materials*. 2012 Nov;12(2):158–164. Available from: http://www.nature.com/doifinder/10.1038/nmat3477.

TS
SR

Liquid Exfoliation of Layered Materials:

Production and Applications of Two-Dimensional Nanosheets

Conor O' Dwyer
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

Layered materials consist of two-dimensional nanosheets which display interesting and useful properties. Liquid exfoliation provides a technique for the production of these nanosheets from a range of layered materials such as graphite, boron nitride and transition metal dichalcogenides, in a scalable and defect-free manner. Various liquid exfoliation techniques exist and are the subject of ongoing research. The combination of scalability and the suspension of dispersed nanosheets in liquid media results in a wide range of promising applications for 2D nanosheets.

Introduction

Layered materials are those which consist of two-dimensional monolayers stacked to form three-dimensional crystals. These structures contain strong in-plane chemical bonds, but weak, usually van der Waals, inter-planar bonds. The most prominent example of a layered material is graphite, which is composed of graphene monolayers, however, numerous other examples exist such as boron nitride (BN), transition metal dichalcogenides (TMDs, e.g. MoS_2 and WS_2), transition metal oxides (TMOs) and layered double hydroxides (LDHs).[1]

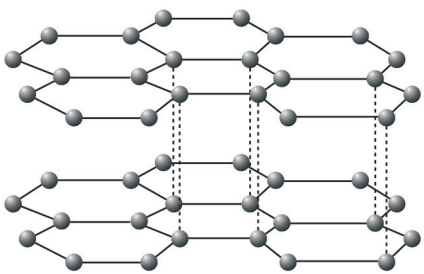


Figure 1A: Layered structure of graphite, which consists of graphene monolayers (hexagonally arranged carbon atoms) *Image adapted from Nicolosi et al.*[1]

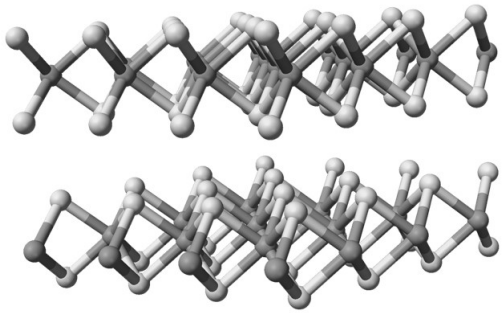


Figure 1B: Layered structure of MoS_2, which consists of a tetragonal-based S and Mo arrangement in each layer. Figures a and b show the general structure for TMDs, with transition metal atoms between layered chalcogen atoms. The tetragonal structure is seen to convert to hexagonal in monolayer form.
Image adapted from Nicolosi et al.[1]

Research into the exfoliation of layered materials was initiated by Geim and Novoselov in 2004 when they demonstrated mechanical cleavage of single-layer graphene from bulk graphite using thin adhesive tape. Exfoliation of monolayer graphene demonstrated unique properties which were not observed in the bulk material.[2] Graphene is transparent due to the atomic thickness of its layers and has a high specific surface area. It is also a zero-gap semiconductor due to the valence and conduction bands meeting at the edges of the Brillouin zone, as a result it demonstrates efficient electrical conductivity properties.[3] These properties showed promise for a wide range of applications, from energy storage,[4] to optoelectronics.[5] The primary disadvantage with the technique used by Geim and Novoselov is that the production rate of 2D crystals is very low. This led to advancements in liquid exfoliation as a method for producing large quantities of 2D nanosheets.[1,6]

Liquid exfoliation techniques

Exfoliation (or delamination) of layered materials in liquids can be separated into two categories, intercalation-based methods, and liquid-phase exfoliation (LPE) methods. Methods for LPE include ultrasonication, ball milling and shear exfoliation.

Intercalation

Intercalation is based on the widening of the interlayer gap between lattice planes of layered crystals through swelling in a solvent and subsequent exfoliation to produce nanosheets. Swelling occurs as a result of the weakening of the van der Waals forces between the layers. The energy required to weaken these interlayer forces is supplied through redox reactions between the intercalating ions and the layered material.[7] When certain metal oxides are introduced to solvents such as water or alcohol, swelling occurs through interaction with solvated ions and allows the layers to be separated through agitation, often using ultrasonication. A good example of this is graphene oxide (GO). Graphite is oxidised with sulfuric

acid and potassium permanganate which results in the addition of hydroxyl and epoxide groups between the layers. Ultrasonication in an appropriate solvent leads to exfoliation to produce GO nanosheets. The oxide layer can then be removed through reduction, but this may result in the formation of defects on the graphene surface which alters the properties substantially from the mechanically cleaved graphene produced by Geim and Novoselov.[8]

Ion Exchange

Ion exchange is a method based on intercalation, where the layered crystals have an exchangeable layer of counterions between the basal planes. Layered double hydroxide (LDH) crystals of the form $M^{2+}_{1-x}M^{3+}_{x}(OH)_2.A^{n-}_{\frac{x}{n}}$ (where M is a first-row transition metal and A is a counterion) are most commonly exfoliated using this technique. The counterions are exchanged for bulky ions in organic solvents which leads to swelling; agitation using ultrasonication or shear mixing then separates the crystals into positively charged nanosheets.[9,10]

Ultrasonication

In intercalation-based methods, ultrasonication is used as an agitation technique to separate layered crystals into nanosheets. It can also be considered as a standalone liquid exfoliation technique. Ultrasonication is the exposure of layered crystals in a solvent to ultrasonic waves that generate cavitation bubbles which collapse, breaking apart the crystals to produce exfoliated nanosheets. The production of liquid-exfoliated monolayer graphene *via* ultrasonication was first shown by the Coleman group in 2008.[11] Other layered nanomaterials such as BN, TMDs and some TMOs have also been delaminated using this method.[12] Ultrasonication presents a number of advantages over previously investigated exfoliation methods. The basal plane of the nanosheets appeared to be defect-free (with only edge defects being observed), thus preserving the properties exhibited by the 2D crystals, and higher concentrations of nanosheets were attained than with previous methods.[13]

The concentration of monolayer nanosheets was seen to be time-dependent; the longer the sonication time, the greater the concentration of monolayers that were obtained. However, the size of the nanosheet flakes was seen to decrease with greater sonication times. Concentrations of 1.2 mg ml^{-1} were reported for sonication times of hundreds of hours.[14] This figure has since increased substantially up to 30 mg ml^{-1} in some cases, but this concentration is still too low to be considered for industrial scale-up.[6] As a result, alternative methods such as ball milling and shear exfoliation have been identified as possible alternatives.

Ball Milling

Ball milling is a more recent liquid-exfoliation technique; it relies on shear and compressive forces to disperse the layered nanomaterials in liquids. The layered material sample is placed in a stirred media mill containing beads of diameter

~100 μm. The grinding action of the beads is responsible for the shear and compressive forces that result in the mechanical delamination of nanosheets.[15] As with ultrasonication, initial research was carried out using graphite. This process again produced monolayer graphene with few defects. It was also shown that 2D crystals of BN and MoS_2 could be produced by the combination of ball milling and ultrasonication.[16] Further research into ball milling is still required, however, the scalability and tunability (with regards to minimising fragmentation and defects) of this method appears promising.

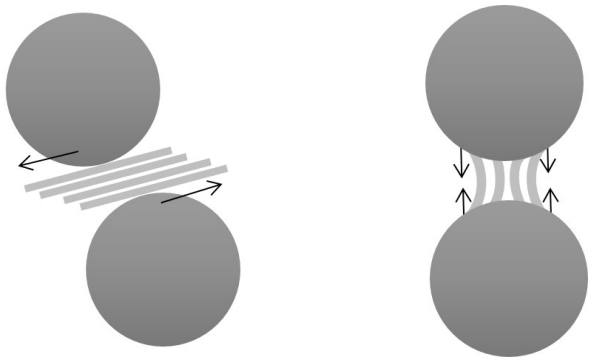


Figure 2: Ball milling for 2D nanosheet production through (a) shear force and (b) compressive force induced exfoliation of layered crystals.

Shear exfoliation

The final LPE technique is shear exfoliation. Initially, shear exfoliation was investigated using rotor/stator mixers to produce graphene nanosheets, along with other layered crystals such as BN and TMDs.[17] This method of exfoliation showed that defect-free few-layer graphene could be produced in a scalable manner as production rates were seen to increase with increasing volume, which is a promising result for industrial scale-up.[6] Nanosheet concentration was also seen to be dependent on the initial concentration of layered material, mixing time and rotor speed; by optimizing each of these parameters, production rates up to 1.3 mg min^{-1} have been achieved.[18] The graphene nanosheets produced by shear exfoliation using rotor/stator mixers are seen to be indistinguishable from those produced by ultrasonication.[17] A disadvantage of rotor/stator mixers is that exfoliation can only occur in close proximity to the mixer, this would prevent scale-up with volume as experimentally shown. However, using rotating blade mixers such as a household blender resolves this problem as it generates shear forces throughout the entire vessel. Although this method is in its infancy, it has already become very popular and demonstrates a scalability for the delamination of many layered nanomaterials.[18]

Stabilization

Liquid exfoliation requires a stabilizer in the form of solvents, surfactants or polymers in order to both increase the ability of the layered materials to be exfoliated and prevent the re-aggregation of exfoliated nanosheets.

Solvents

Exfoliated nanosheets are stabilized in liquids when the net energy cost of mixing for the solute and solvent is minimised. The most important parameter for stabilization, and thus high nanosheet concentrations, is surface tension, γ. For graphene and WS_2 it was seen that $\gamma \sim 40 mJ\ m^{-2}$ produced the highest nanosheet concentrations.[12] Not only do suitable solvents stabilize the 2D crystals in solution, they also prevent the re-aggregation by means of slowing down reaction rates on the nanosheet surfaces.[6] The primary problem with solvents suitable for exfoliation is their toxicity and incompatibility with further processing techniques. N-Methylpyrrolidone (NMP), for example, is the most widely used solvent for LPE, however, it has a high boiling point which makes it difficult to remove after exfoliation and has also been shown to be teratogenic.[6] As such, attempts have been made to find alternative solvents. Unfortunately, water has a surface tension of $72 mJ\ m^{-2}$ and so is not suitable for exfoliation of layered crystals such as graphite or WS_2.[1] Low boiling point solvents such as ethanol and iso-propyl alcohol (IPA) were also investigated, however, with $\gamma \sim 25 mJ\ m^{-2}$ for these solvents, the concentration of nanosheets obtained was too low.[6] Attempts were therefore made to use a mixture of water/ethanol and water/IPA in order to produce an effective stabilizer with a low boiling point and non-toxicity. Despite the ability to tune γ based on water/alcohol composition, the concentration of exfoliated sheets was still lower when compared with NMP. Furthermore, the stability of the mixtures with respect to evaporation and variation in temperature is a major problem.[19]

Surfactants

Stabilization *via* surfactant molecules, such as sodium dodecylbenzenesulfonate (SDBS) or sodium cholate (SC), in water provide an alternative to solvents. When layered materials are exfoliated in a surfactant/water environment, the nanosheets become coated with surfactant molecules through a non-covalent interaction between the tail group of the surfactant molecule and the basal plane of the nanosheet. The head group interacts with the water and results in stabilization of the delaminated layers and prevents re-aggregation through steric or electrostatic repulsion.[20]

Polymers

Polymers can be used to stabilize exfoliated graphene and other 2D nanosheets in solvents *via* steric repulsion between polymers which have adsorbed onto the layers. By determining the balance between adsorption energy and loop/tail

entropy of the polymer in solution, predictions about the suitability of the solvent/polymer pair can be made.[21] A disadvantage of polymer stabilization is the lower concentration of monolayer crystals produced compared with either solvent or surfactant stabilization. Additionally, polymers can be difficult to remove after exfoliation.[22] However, in certain cases, monolayers which are enriched with polymers can retain their properties; for example, WS_2 with adsorbed poly(vinyl alcohol) still exhibits the optoelectronic properties observed in pristine WS_2.[23]

Characterisation

Evidence for the exfoliation of layered materials to produce nanosheets can be determined by transmission electron microscopy (TEM). A small amount of the crystal dispersions are dropped onto a holey carbon grid and images such as those shown in Fig. 3a and 3b are obtained. It is clear from these images that 2D crystals are obtained. Fragments with thickness of more than a few layers are rarely observed after sufficient exfoliation. Fig. 3a and Fig. 3b show graphene mono and multilayers respectively. By counting the number of layers present in a dispersion of 100 flakes, for example, a distribution of the number of layers per flake can be determined. In most cases, less than 10 layers are observed per flake, with the distribution concentrated in the region of <5 layers, this further provides evidence of exfoliation.[11]

AFM and SEM are similar techniques to TEM for imaging nanosheet flakes; the liquid suspended flakes are deposited on a membrane suitable to both the imaging technique and material in question. The size and height of the nanosheets can be determined from the images.

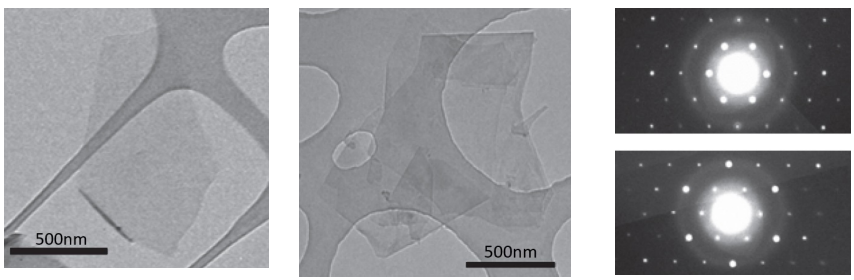


Figure 3: (A) and (B) show TEM images of a liquid exfoliated graphene mono and multilayers respectively deposited on a holey carbon grid. (C) shows the electron diffraction patterns for single (top) and double (bottom) layers of graphene
Images courtesy of Hernandez et al.[11]

Monolayers can be distinguished from multilayers by electron diffraction (ED). A hexagonal pattern is produced in both cases in Fig. 3c as expected for graphene. Evidence of a monolayer is provided by more intense inner spots compared with more intense outer spots for mulitlayer graphene.[11,12]

Raman spectroscopy is used to confirm that the delaminated 2D crystals are defect free. Raman spectra of exfoliated flakes can be compared with the bulk material to show that the flakes have the expected structure; defects could alter the structure or introduce peaks to the spectrum which are not present in that for the bulk crystal. Furthermore, X-ray photoelectron spectroscopy (XPS) and infrared (IR) spectroscopy can by used to confirm the defect-free nature of the nanosheets. In XPS, defect-free graphene, for example, can be indicated if the spectrum contains points corresponding to only graphitic carbon and residual solvent. This demonstrates the absence of oxidation which is the primary defect associated with GO. Similarly, the absence of oxidation groups or other defects in the IR spectra of layered materials indicates its defect-free nature.[11,12]

Applications

Liquid exfoliation provides a useful and scalable route for the production of 2D nanomaterials. The production of these nanosheets as suspensions in liquids provides applications and processing techniques that would otherwise be difficult. Liquid exfoliation allows the deposition of 2D crystals onto surfaces to form thin-films through techniques such as inkjet printing. Some applications of nanosheets include electrodes for energy storage devices such as supercapacitors, electronics and optoelectronics, and composities.[1,24]

Supercapacitor electrodes can be coated with 2D nanomaterials such as graphene in order to greatly increase the energy storage capacity due to the large surface area of graphene; surface areas of up to 2000 m^2g^{-1} are possible.[4] Graphene-based supercapacitors have yielded energy densities close to those of lead-acid batteries. A number of other 2D nanomaterials are also suitable for energy storage, namely TMDs, and can introduce varying electrochemical properties.[24] Therefore, combinations of nanosheets may lead to optimal energy storage electrodes.[1]

Liquid exfoliation allows composites of nanosheets with polymers and other nanostructures to be formed, which can lead to interesting material properties. For example, WS_2/carbon nanotube composites display the characteristics of WS_2 films with the added conductivity from carbon nanotubes. Composites such as these have applications for photovoltaic devices and supercapacitor electrodes.[12]

As a result of the mostly defect-free nanosheets produced from liquid exfoliation, the 2D materials retain their electronic structure and properties. This allows for a wide range of applications in electronics and optoelectronics. For example, TMDs

have tunable bandgaps depending on the number of nanosheet layers, which could prove useful for applications such as transistors, photodetectors and electroluminescent devices.[25] Additionally, TMDs such as MoS_2 show transistor stability under mechanical deformation which provides interesting applications for flexible and transparent electronics and optoelectronics.[1]

Conclusions

A number of liquid exfoliation techniques have been presented, with possibly the most promising method for future research and scalability to industry being shear exfoliation. A wide range of layered materials such as graphite, BN and TMDs have already been delaminated using this technique, however, many still remain for future investigation. Nanosheets display interesting and useful properties which are retained after liquid exfoliation, this along with the increased processing ability of the 2D crystals in liquids allows for many applications in areas such as electronics and optoelectronics, energy storage, and photovoltaics. The research into liquid exfoliation is still young, however, there is no doubt that there is much promising research to come in the future.

Acknowledgements

The author would like to thank Mr. I. Harley, Mr. J. Lyons, Mr. J. Boland and Prof. I. Shvets for their helpful input and appraisal.

References

1. Nicolosi, V.; Chhowalla, M.; Kanatzidis, M. G.; Strano, M. S.; Coleman, J. N. Liquid Exfoliation of Layered Materials. *Science* (80- .). 2013, 340 .

2. Geim, A. K. Graphene: Status and Prospects. *Science* (80-.). 2009, 324, 1530–1534.

3. Geim, A. K.; Novoselov, K. S. The rise of graphene. *Nat. Mater.* 2007, 6, 183–191.

4. Zhu, Y.; Murali, S.; Stoller, M. D.; Ganesh, K. J.; Cai, W.; Ferreira, P. J.; Pirkle, A.; Wallace, R. M.; Cychosz, K. A.; Thommes, M.; Su, D.; Stach, E. A.; Ruoff, R. S. Carbon-based supercapacitors produced by activation of graphene. *Science* (80-.). 2011, 332, 1537–1541.

5. Schwierz, F. Graphene transistors. *Nat. Nanotechnol*. 2010, 5, 487–496.

6. Bonaccorso, F.; Bartolotta, A.; Coleman, J. N.; Backes, C. 2D-Crystal-Based Functional Inks. *Adv. Mater*. 2016, 28, 6136–6166.

7. Jacobson, A. J.; Nazar, L. F.; Jacobson, A. J.; Nazar, L. F. Encycl. Inorg.Bioinorg. Chem.; *John Wiley & Sons, Ltd: Chichester,* UK, 2011.

8. Dreyer, D. R.; Park, S.; Bielawski, C. W.; Ruoff, R. S. The chemistry of graphene oxide. *Chem. Soc. Rev.* 2010, 39, 228–240.

9. Ma, R.; Liu, Z.; Li, L.; Iyi, N. Exfoliating layered double hydroxides in formamide: a method to obtain positively charged nanosheets. *J. Mater. Chem*. 2006, 16, 3809.

10. Ma, R.; Sasaki, T. Nanosheets of Oxides and Hydroxides: Ultimate 2D Charge-Bearing Functional Crystallites. *Adv. Mater.* 2010, 22, 5082–5104.

11. Hernandez, Y. *et al*. High-yield production of graphene by liquid-phase exfoliation of graphite. *Nat. Nanotechnol.* 2008, 3, 563–568.

12. Coleman, J. N. *et al*. Two-Dimensional Nanosheets Produced by Liquid Exfoliation of Layered Materials. *Science* (80-.). 2011, 331 .

13. Coleman, J. N. Liquid Exfoliation of Defect-Free Graphene. *Acc. Chem. Res.* 2013, 46, 14–22.

14. Khan, U.; O'Neill, A.; Lotya, M.; De, S.; Coleman, J. N. High-Concentration Solvent Exfoliation of Graphene. *Small* 2010, 6, 864–871.

15. Knieke, C.; Berger, A.; Voigt, M.; Taylor, R. N. K.; Rohrl, J.; Peukert, W. Scalable production of graphene sheets by mechanical delamination. *Carbon N.Y.* 2010, 48, 3196–3204.

16. Yao, Y.; Lin, Z.; Li, Z.; Song, X.; Moon, K.-S. Large-scale production of two-dimensional nanosheets. *J. Mater. Chem.* 2012, 22, 13494.

17. Paton, K. R. *et al*. Scalable production of large quantities of defect-free few-layer graphene by shear exfoliation in liquids. *Nat. Mater.* 2014, 13, 624–630.

18. Varrla, E.; Backes, C.; Paton, K. R.; Harvey, A.; Gholamvand, Z.; McCauley, J.; Coleman, J. N. Large-Scale Production of Size-Controlled MoS_2 Nanosheets by Shear Exfoliation. *Chem. Mater*. 2015, 27, 1129–1139.

19. Halim, U.; Zheng, C. R.; Chen, Y.; Lin, Z.; Jiang, S.; Cheng, R.; Huang, Y.; Duan, X. A rational design of cosolvent exfoliation of layered materials by directly probing liquid–solid interaction. *Nat. Commun.* 2013, 4, 2213.

20. Lotya, M.; Hernandez, Y.; King, P. J.; Smith, R. J.; Nicolosi, V.; Karlsson, L. S.; Blighe, F. M.; De, S.; Wang, Z.; McGovern, I. T.; Duesberg, G. S.; Coleman, J. N. Liquid Phase Production of Graphene by Exfoliation of Graphite in Surfactant/ Water Solutions. *J. Am. Chem. Soc.* 2009, 131, 3611–3620.

21. May, P.; Khan, U.; Hughes, J. M.; Coleman, J. N. Role of Solubility Parameters in Understanding the Steric Stabilization of Exfoliated Two-Dimensional Nanosheets by Adsorbed Polymers. *J. Phys. Chem*. C 2012, 116, 11393–11400.

22. Ferrari, A. C. *et al*. Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems. *Nanoscale* 2015, 7, 4598–4810.

23. Vega-Mayoral, V.; Backes, C.; Hanlon, D.; Khan, U.; Gholamvand, Z.; O'Brien, M.; Duesberg, G. S.; Gadermaier, C.; Coleman, J. N. Photoluminescence from Liquid-Exfoliated WS2 Monomers in Poly(Vinyl Alcohol) Polymer Composites. *Adv. Funct. Mater.* 2016, 26, 1028–1039.

24. Chhowalla, M.; Shin, H. S.; Eda, G.; Li, L.-J.; Loh, K. P.; Zhang, H. The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets. *Nat. Chem.* 2013, 5, 263-275.

25. Wang, Q. H.; Kalantar-Zadeh, K.; Kis, A.; Coleman, J. N.; Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. Nat. Nanotechnol. 2012, 7, 699–712.10. Ma, R.; Liu, Z.; Li, L.; Iyi, N. *J. Mater. Chem*. 2006, 16, 3809.

11. Ma, R.; Sasaki, T. *Adv. Mater.* 2010, 22, 5082–5104.

12. Hernandez, Y. *et al*. *Nat. Nanotechnol.* 2008, 3, 563–568.

13. Coleman, J. N. *et al*. *Science* (80-.). 2011, 331, year.

14. Coleman, J. N. *Acc. Chem. Res*. 2013, 46, 14–22.

15. Khan, U.; O'Neill, A.; Lotya, M.; De, S.; Coleman, J. N. *Small* 2010, 6, 864–871.

16. Knieke, C.; Berger, A.; Voigt, M.; Taylor, R. N. K.; Rohrl, J.; Peukert, W. Carbon N. Y. 2010, 48, 3196–3204.

17. Yao, Y.; Lin, Z.; Li, Z.; Song, X.; Moon, K.-S. *J. Mater. Chem.* 2012, 22, 13494.

18. Paton, K. R. *et al*. *Nat. Mater.* 2014, 13, 624–630.

19. Varrla, E.; Backes, C.; Paton, K. R.; Harvey, A.; Gholamvand, Z.; McCauley, J.; Coleman, *J. N. Chem. Mater.* 2015, 27, 1129–1139.

20. Halim, U.; Zheng, C. R.; Chen, Y.; Lin, Z.; Jiang, S.; Cheng, R.; Huang, Y.; Duan, X. *Nat. Commun.* 2013, 4, 2213.

21. Lotya, M.; Hernandez, Y.; King, P. J.; Smith, R. J.; Nicolosi, V.; Karlsson, L. S.; Blighe, F. M.; De, S.; Wang, Z.; McGovern, I. T.; Duesberg, G. S.; Coleman, *J. N. J. Am. Chem.* Soc. 2009, 131, 3611–3620.

22. May, P.; Khan, U.; Hughes, J. M.; Coleman, *J. N. J. Phys. Chem.* C 2012, 116, 11393–11400.

23. Ferrari, A. C. *et al*. *Nanoscale* 2015, 7, 4598–4810.

24. Vega-Mayoral, V.; Backes, C.; Hanlon, D.; Khan, U.; Gholamvand, Z.; O'Brien, M.; Duesberg, G. S.; Gadermaier, C.; Coleman, *J. N. Adv. Funct. Mater.* 2016, 26, 1028–1039.

25. Chhowalla, M.; Shin, H. S.; Eda, G.; Li, L.-J.; Loh, K. P.; Zhang, *H. Nat. Chem.* 2013, 5, 263–275.

26. Wang, Q. H.; Kalantar-Zadeh, K.; Kis, A.; Coleman, J. N.; Strano, M. S. *Nat. Nanotechnol.* 2012, 7, 699–712.

Designing Artificial Photosynthetic Systems using Nanomaterials for Energy Production and Storage

Aoife Gregg
Junior Sophister
Nanoscience, Physics & Chemistry of Advanced Materials

Artificial photosynthesis utilises light to produce hydrogen by splitting water into its constituent elements. This direct conversion of solar energy to hydrogen fuel simultaneously produces and stores energy. Hydrogen provides a viable alternative to fossil fuels, and artificial photosynthetic methods allow it to be produced sustainably. This review discusses the main approach to artificial photosynthesis: photoelectrochemical cells, which use semiconductors to drive the electrolysis of water. The use of nanomaterials to improve these systems is essential to their future real-life implementation, as nanomaterials have tunable electrical and chemical properties, can have high light-capture and conversion efficiencies, and are often composed of earth-abundant materials. The key nanomaterials investigated for this purpose are discussed, including graphene-derived materials, silver nanowires, and a variety of nanosheets.

Introduction

New challenges in energy production and environmental preservation mean that reliance on fossil fuels is no longer feasible. Energy demands are set to double in the next 30 years.[1,2] The efficient production of hydrogen is essential for potentially important future technologies, such as hydrogen fuel cells and the synthesis of fuels from syngas and CO_2.[3] Hydrogen fuel cells emit only water and are already being tested for applications in hydrogen-powered transport.[4] H_2 has the highest energy density of chemical fuels (140 MJ kg^{-1}) on a gravimetric basis, which is higher than most hydrocarbon fuels (40 – 50 MJ kg^{-1}).[5] Currently, hydrogen is produced in industry by the fossil fuel-based steam reforming reaction, which uses hydrocarbons and produces carbon dioxide as a by-product.[6] Artificial photosynthesis poses a sustainable and environmentally friendly means of capturing solar energy and storing it in the form of hydrogen.

Artificial photosynthesis describes the method of capturing light energy and using it to split water into hydrogen and oxygen. This stores the energy in the form of hydrogen and therefore compensates for the intermittency of sunlight as a form of power. This can be executed by using semiconductors as photocatalysts, which are materials in which light energy drives the catalysis of a reaction. The absorption of light in a semiconductor excites an electron from the valence band to the conduction band, producing an electron-hole pair. The electron and hole can then initiate redox reactions, which are chemical reactions that involve the transfer of electrons.[7]

Since the seminal paper by Fujishima and Honda in 1972,[8] semiconducting photocatalysts have been used as electrodes in photoelectrochemical cells. In these cells, connected electrodes (at least one of which is a photocatalyst) are immersed in an electrolyte – in this case, water. Light-generated electron-hole pairs are separated and reach the electrolyte by travelling either to the surface of the semiconductor or through the wires to a counter electrode. Once they reach the water, holes oxidise the water to produce oxygen (the oxygen evolution reaction – OER) and electrons reduce the water to produce hydrogen (the hydrogen evolution reaction – HER).[9–11]

It has been shown that photoelectrochemical cells have a theoretical efficiency limit of roughly 40% for dual-junction devices,[12] and experimental efficiencies of nearly 20% have been achieved.[13] The use of nanomaterials improves photoelectrochemical cells. They can be used to optimise the morphology of a photoelectrochemical cell because they have a high aspect ratio. They can be used as semiconducting photoelectrode catalysts, or as co-catalysts to improve the kinetics of the reaction. Nanomaterials can replace expensive, rare materials currently needed as catalysts, such as platinum and gold.[7,9] They allow for flexibility,[10,14] greater tunability,[15] greater light absorption efficiency,[16] and greater reaction efficiency.[7,15,17] Owing to the great number of nanomaterials designed for use in artificial photosynthesis, it not possible to discuss them all in this review, so representative samples which show great potential have been chosen.

Semiconductors as Photocatalysts

When a semiconducting material is irradiated, electron-hole pairs will be generated by photons of light with energy greater than or equal to the band gap of the material. If no action is taken to separate them, these electron-hole pairs are likely to recombine and release thermal energy. The pairs must be separated for photocatalysis to occur.

If the semiconductor is placed in a solution, charge transfer spontaneously occurs at the interface if there is a difference in tendency to lose or gain electrons of the two phases i.e. the Fermi energy of the solid and the redox potential of the solution.[11] This charge flow continues until the system reaches equilibrium, which occurs when the Fermi energy of the electrons in the semiconductor is equal to the redox

potential of the solution.[10] This produces a space-charge layer: a region on either side of the interface where the electric charge distribution differs from the bulk. On the semiconductor side, the valence and conducting band potentials near the surface of the semiconductor are altered from the values in the bulk semiconductor. When a positive charge accumulates at the surface of the semiconductor, the band potentials are decreased (Figure 1b). When a negative charge accumulates at the surface, the band potentials increase (Figure 1c).

The resulting charge distribution creates an electric field which splits the electron-hole pairs before they can recombine. The separated electron and hole can each then catalyse redox reactions by interacting with the molecules adsorbed on the working electrodes.[18] A redox reaction may be catalysed by the photogenerated electron and hole when the standard redox potentials lie between the potential of the conduction band and the valence band.[17,19]

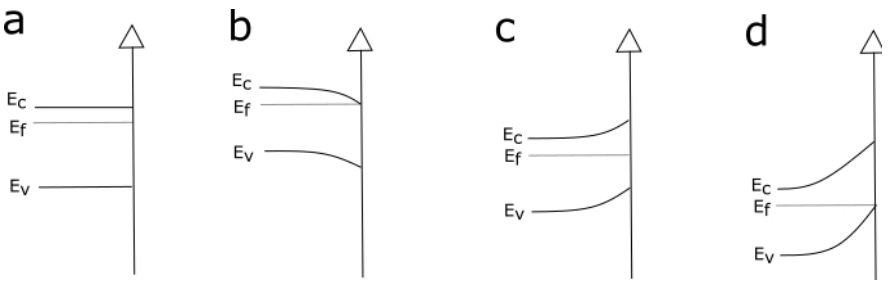


Figure 1: Band bending for an n-type semiconductor-electrolyte interface, where the semiconductor is to the left of the arrow and the electrolyte is to the right. E_c = conduction band potential, E_v = valence band potential, E_f = Fermi energy.
(A) flat band potential, where no space-charge layer exists as the Fermi energy of the semiconductor is equal to the redox potential of the electrolyte;
(B) accumulation layer, where electrons have travelled into the bulk of the semiconductor and holes have gathered on the interface;
(C) depletion layer, where electrons have moved from the semiconductor to the electrolyte;
(D) inversion layer, where the number of electrons in the semiconductor has been reduced to below its intrinsic level, generating a p-type semiconductor at the interface.[10,18]

Photoelectrochemical Cells

In a basic photoelectrochemical cell, an n-type semiconductor is immersed in water. Holes move towards the interface and oxidise water to oxygen by combining with an electron donated by the water. The electrons pass through a circuit to an inert counter electrode also immersed in water, where they reduce the water to hydrogen. This leads to the overall water-splitting reaction:[10,11]

$$2H_2O \rightarrow O_2 + 2H_2$$

To improve the catalytic ability of the cell, semiconductors can be used for both electrodes and cocatalysts can be used in conjunction with the semiconductors. Two practical requirements must be met when designing the morphology of a photoelectrochemical cell for splitting water. A robust mechanism for the separation of products is vital, as a $H_{2(g)}$-$O_{2(g)}$ mixture can be explosive. The combination of a hydrogen evolution reaction (HER) which consumes protons at one electrode and an oxygen evolution reaction (OER) which generates protons at the other electrode leads to the development of a significant pH gradient. This reduces the efficiency of the cell by requiring more energy to drive the reaction, so strategies are needed to avoid this gradient.[7]

A dual microwire array (Figure 2) has many features which serve to optimise the photolysis of water.[7] The membrane is impermeable to H_2 and O_2 but allows the transfer of ions, which separates the products to ensure safe operation and allows for the neutralisation of the pH gradient. It is transparent and it allows for electrical transport between the microwire arrays.[20] The high aspect ratio of the semiconducting microwires provides a short path along the radial dimension across which minority carriers can diffuse for charge collection, while maintaining a high surface area for the absorption of light and for the semiconductor-water interface. The microwires are decorated with co-catalysts. Different semiconductors are used for the photoanode and the photocathode so that the photoanode has a larger band gap energy than the photocathode. This allows the dual microwire array to act as a tandem photovoltaic cell. Upon irradiation, high energy light is absorbed by the top semiconductor, which has a large band gap.[7,20] Light with an energy lower than the band gap energy of the top semiconductor passes through

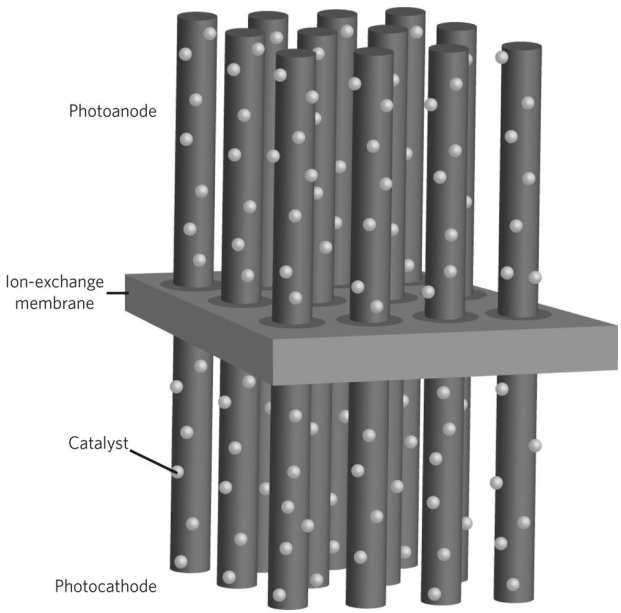


Figure 2: Dual microwire array. A tandem cell composed of a photoanode (which absorbs blue light and effects water oxidation), a photocathode (which absorbs red light and effects water reduction), and a polymeric membrane which separates the products and allows for the transfer of ions. The photoelectrodes are decorated with electrocatalysts. *Reprinted by permission from Macmillan Publishers Ltd: Nature Nanotechnology*[7]*, copyright (2016).*

the top semiconductor and is absorbed by the bottom semiconductor, which has a smaller band-gap energy. The use of multiple junctions overcomes the theoretical efficiency limit of a single-junction cell, which is about 33.16%.[21]

Nanomaterials for Photoelectrochemical Cells

TiO_2 has long been used in photosynthetic cells for water splitting.[8,10] However, the large band gap of TiO_2 means that it only absorbs the ultraviolet part of the solar spectrum, which severely limits its efficiency as a photocatalyst as ultraviolet light makes up only 4% of the incoming solar energy on the Earth's surface.[22] Nanomaterials hold great potential to optimise the absorption of light, as well as the catalytic activity of the semiconductors and cocatalysts that compose a photoelectrochemical cell. This could allow artificial photosynthesis to become a viable option for energy production.

The use of multi-junction photovoltaic cells provides a mechanism by which light absorption can be increased. Theoretically, an infinite number of junctions under concentrated sunlight would have a limiting efficiency of 86.8%.[23] A multi-junction photovoltaic cell relies on the passage of light through the multiple junctions, which means that the conductors used to collect charge at each junction must be transparent. Graphene and graphene-based materials have been used to produce flexible, cheap, and optically transparent conductors.[14,24] Graphene has high chemical and thermal stability, low sheet resistance and is highly optically transparent. However, current synthesis mechanisms for large-area monolayer graphene can result in many topological defects, which results in highly resistive grain boundaries that trap charge carriers.[14] It has been found that the combination of graphene with silver nanowires (AgNWs) eliminates the line defects and disruptions in graphene while maintaining high optical transparency.[14,25] AgNWs are used because of their high conductivity, flexibility, and facile fabrication. However, the AgNWs are unstable against oxidation when exposed to air and water. Encapsulating the AgNWs in a transparent polymer (such as PVA) combats this. A 'sandwich' structure has been produced: a layer of graphene was spray-coated with AgNWs, then spin-coated with PVA-124.[14] The presence of the PVA gave the electrode long-term stability by inhibiting the oxidation of the AgNWs and by effecting more intimate contact and better adhesion between the graphene and the AgNWs. This resulted in a further decrease in sheet resistance. Future research may improve these properties further by using multifunctional polymers.[14]

Many nanomaterials have been designed to act as photocatalysts for hydrogen generation. Sulfide-based bionanocomposite films, nanospheres and nanorods have all been used for this purpose.[17] Photoelectrocatalysis using nanosheets as electrodes has shown to be particularly effective. Structural distortions in the nanosheets lead to an increase in the density of states of the valence band, which improves optical absorption and charge carrier mobility. The ultrathin and high aspect ratio nature of the sheets shortens the diffusion length for the charge carriers, which enables them to reach the surface and take part in chemical reactions

before recombination.[15] Implementing truly atomistic thickness in a nanosheet would fundamentally eliminate bulk electron-hole recombination. With this goal in mind, the synthesis and use of nanosheets is challenging. Tin sulfide (SnS) semiconductor nanosheets with two-atom thickness have been synthesised by exfoliating bulk SnS. These nanosheets have shown photon-to-current conversion efficiency of 67.1% at 490 nm, which is significantly higher than efficiencies of other visible-light-driven water splitting.[16] SnS has many properties ideal for the synthesis of nanosheets: it uses earth-abundant materials, bulk SnS is composed of layers held together only by van der Waals interactions, the monolayers are held together strongly by chemical bonds, and the monolayer is composed of 100% exposed surface atoms.

Co-catalysts are used to improve the efficiency of semiconductor catalysts by reducing the photogenerated voltage needed to drive the reaction or by increasing the photogenerated voltage. Nanomaterials are being investigated to replace co-catalysts such as platinum, which are expensive, unsustainable materials that limit the scalability of artificial photosynthetic systems. In particular, transition-metal-dichalcogenide (TMD) nanosheets have been used because of the highly active chalcogenide atoms at their edges, their tunable band gap, and their large specific surface area. TMD nanosheets can be loaded onto semiconductors, and these composite materials can have an improved electron-hole separation and a greater range of light absorption.[9] MoS_2 (a TMD) nanosheets have been widely used; CdS loaded with only 0.2 wt% MoS_2 had a photocatalytic activity 36 times greater than the photocatalytic activity of CdS, and TiO_2-nanowire/MoS_2-nanosheet hybrid nanostructures had a hydrogen generation rate 83 times greater than pure TiO_2 nanowires.[9,26] The effectiveness of the nanosheets can be further improved by modifications to the basic nanosheet structure, such as doping.[27] Along with their use in enhancing photocatalysis, TMD and other nanosheets can be used as electrocatalysts, which work as cocatalysts in conjunction with photocatalysts to improve the kinetics of a photoelectrochemical cell.[15,28] Other nanomaterials can also act as effective electrocatalysts: carbon nanotubes can be used to encapsulate non-precious metal catalysts, which are otherwise unstable in acidic electrolytes.[3]

Large-scale production of nanosheets is a major area of research. Mechanisms such as chemical vapour deposition (CVD) and liquid-based exfoliation have been shown to produce high-quality graphene films. CVD can produce highly uniform, large area monolayer graphene. However, it is an expensive process and unlikely to be able to meet the demands of commercial applications.[29,30] Liquid-based exfoliation can be used to produce industrial-scale amounts of defect-free nanosheets from a range of layered materials, such as graphene and TMDs.[31] However, current issues such as low yield, low monolayer content and broad distributions of thickness and area limit its applications.[30]

Advances in the field of solar cells are directly applicable to artificial photosynthesis. Dye-sensitised solar cells (DSSCs) and perovskite solar cells have shown huge progress in recent years. The former uses dyes to achieve a high quantum

efficiency.[32] The latter have low production costs and are the fastest-improving solar technology to date, showing an increase from 3.8% to 22.1% between 2009 and 2016.[33] Implementing these new technologies for use in hydrogen production could greatly improve the solar conversion efficiency of the water-splitting photoelectrochemical system.

Powder Photocatalysis

While direct photocatalysis systems currently show efficiencies lower than those of photoelectrochemical systems, they are simple constructions which have the potential for cheap, large-scale production of H_2. This is because they are typically composed just of photocatalytic powders suspended in water.[34] Nitrogen-doped graphene oxide quantum dots (NGO-QDs) have been used for this purpose. They have a large accessible surface area in aqueous solutions, and their conduction and valence band potentials are suitable for H_2 and O_2 evolution. Pure graphene is a zero-band gap semiconductor as its π and π^* orbitals touch at the Brillouin zone, so its electronic properties can be tuned by chemical modification and size changes.[35] The formation of C-O bonds in graphene when synthesising graphene oxide (GO) alters the original orbitals and confines the π electrons, thus creating a band gap. The small size of the NGO-QDs introduces quantum confinement, which separates the π and π^* orbitals and creates a band gap. As oxygen is more electronegative than carbon, GO is a p-doped semiconductor and so an accumulation layer is formed when it is immersed in water, which catalyses the reduction of water to hydrogen. Doping GO with nitrogen at the edges transforms it into an n-type semiconductor at those locations, which catalyses the oxidation of water to oxygen. This creates a visible-light-stimulated photocatalyst that catalyses both parts of the water-splitting reaction. Earth-abundant materials are used, so these NGO-QDs provide a viable option for sustainable, environmentally friendly hydrogen production.[36]

Discussion and Conclusions

Recent research has shown that artificial photosynthesis may be a solution to the challenges of energy production and storage. Nanomaterials have shown to be key in making these artificial photosynthetic systems a viable option; the efficiency, low cost and sustainability that they lend to the artificial photosynthetic systems will be needed for real-life implementation.

Future research must further increase the efficiencies of hydrogen production systems by developing precise control of the electronic and structural properties of nanomaterials to maximise light absorption, reduce charge carrier recombination, and improve the kinetics of the HER and OER. Current and future nanomaterials that have only been used in isolation should be incorporated into full photosynthetic

systems. The main challenge facing the use of nanomaterials, and therefore the efficient production of hydrogen, is mass production.[15] Although high-quality, large-scale production methods do exist, they do not yet reach the standards needed for commercial applications.

Acknowledgements

The author would like to thank Mr. J. Lyons, Mr. A. Kelly and Prof. I. Shvets for their helpful input and appraisal.

References

1. Dai L, Chang DW, Baek J-B, Lu W. Carbon nanomaterials for advanced energy conversion and storage. *Small* 2012; 8: 1130–66.

2. Gong Y, Wang J, Wei Z, Zhang P, Li H, Wang Y. Combination of carbon nitride and carbon nanotubes: Synergistic catalysts for energy conversion. *ChemSusChem* 2014; 7: 2303–2309.

3. Deng J, Ren P, Deng D, Yu L, Yang F, Bao X *et al*. Highly active and durable non-precious-metal catalysts encapsulated in carbon nanotubes for hydrogen evolution reaction. *Energy Environ Sci* 2014; 7: 1919.

4. McCafferty G. Zero emission hydrogen train to run in Germany. *CNN*. 2016. http://edition.cnn.com/2016/11/03/europe/germany-zero-emissions-train/

5. Ong W-J, Tan L-L, Ng YH, Yong S-T, Chai S-P. Graphitic Carbon Nitride (g-C 3 N 4)-Based Photocatalysts for Artificial Photosynthesis and Environmental Remediation: Are We a Step Closer To Achieving Sustainability? *Chem Rev* 2016; 116: 7159–7329.

6. Breen JP, Burch R, Coleman HM. Metal-catalysed steam reforming of ethanol in the production of hydrogen for fuel cell applications. *Appl Catal B Environ* 2002; 39: 65–74.

7. Lewis NS. Developing a scalable artificial photosynthesis technology through nanomaterials by design. *Nat Nanotechnol* 2016; 11: 1010–1019.

8. Fujishima A, Honda K. Electrochemical photolysis of water at a semiconductor electrode. *Nature* 1972; 238: 37–8.

9. Lu Q, Yu Y, Ma Q, Chen B, Zhang H. 2D Transition-Metal-Dichalcogenide-Nanosheet-Based Composites for Photocatalytic and Electrocatalytic Hydrogen Evolution Reactions. *Adv. Mater.* 2016; 28: 1917–1933.

10. Grätzel M. Photoelectrochemical cells. *Nature* 2001; 414: 338–344.

11. Bard AJ. Photoelectrochemistry. *Science* (80-) 1980; 207.

12. Fountaine KT, Lewerenz HJ, Atwater HA. Efficiency limits for photoelectrochemical water-splitting. *Nat Commun* 2016; 7: 13706.

13. Licht S, Wang B, Mukerji S, Soga T, Umeno M, Tributsch H. Efficient Solar Water Splitting, Exemplified by RuO 2 -Catalyzed AlGaAs/Si Photoelectrolysis. *J Phys Chem B* 2000; 104: 8920–8924.

14. Zhang Q, Di Y, Huard CM, Guo LJ, Wei J, Guo J *et al*. Highly stable and stretchable graphene–polymer processed silver nanowires hybrid electrodes for flexible displays. *J Mater Chem C* 2015; 3: 1528–1536.

15. Deng D, Novoselov KS, Fu Q, Zheng N, Tian Z, Bao X. Catalysis with two-dimensional materials and their heterostructures. *Nat Nanotechnol* 2016; 11: 218–230.

16. Sun Y, Sun Z, Gao S, Cheng H, Liu Q, Lei F *et al*. All-Surface-Atomic-Metal Chalcogenide Sheets for High-Efficiency Visible-Light Photoelectrochemical Water Splitting. *Adv Energy Mater* 2014; 4: 1300611.

17. Li X, Yu J, Jaroniec M, Li X, Wen J, Low J *et al*. Hierarchical photocatalysts. *Chem Soc Rev* 2016; 45: 2603–2636.

18. Linsebigler AL, Lu G, Yates JT. Photocatalysis on TiO2 Surfaces: Principles, Mechanisms, and Selected Results. *Chem Rev* 1995; 95: 735–758.

19. Xu Y, Schoonen MAA. The absolute energy positions of conduction and valence bands of selected semiconducting minerals. *Am Mineral* 2000; 85.

20. Spurgeon JM, Walter MG, Zhou J, Kohl PA, Lewis NS, Johnson EL *et al*. Electrical conductivity, ionic conductivity, optical absorption, and gas separation properties of ionically conductive polymer membranes embedded with Si microwire arrays. *Energy Environ Sci* 2011; 4: 1772.

21. Rühle S. Tabulated values of the Shockley–Queisser limit for single junction solar cells. *Sol Energy* 2016; 130: 139–147.

22. Li Z, Luo W, Zhang M, Feng J, Zou Z. Photoelectrochemical cells for solar hydrogen production: current state of promising photoelectrodes, methods to improve their properties, and outlook. *Energy Environ Sci* 2013; 6: 347–370.

23. Green M. Tandem Cells. In: Third Generation Photovoltaics. *Springer Berlin Heidelberg*, 2003, p 65.

24. Xu Y, Liu J. Graphene as Transparent Electrodes: Fabrication and New Emerging Applications. *Small.* 2016; 12: 1400–1419.

25. Kholmanov IN, Magnuson CW, Aliev AE, Li H, Zhang B, Suk JW *et al*. Improved electrical conductivity of graphene films integrated with metal nanowires. *Nano Lett* 2012; 12: 5679–5683.

26. Shen M, Yan Z, Yang L, Du P, Zhang J, Xiang B. MoS2 nanosheet/TiO2 nanowire hybrid nanostructures for enhanced visible-light photocatalytic activities. *Chem Commun* 2014; 50: 15447–15449.

27. Deng J, Li H, Xiao J, Tu Y, Deng D, Yang H *et al*. Triggering the electrocatalytic hydrogen evolution activity of the inert two-dimensional MoS_2 surface *via* single-atom metal doping. *Energy Environ Sci* 2015; 8: 1594–1601.

28. Liu Y, Cheng H, Lyu M, Fan S, Liu Q, Zhang W *et al*. Low overpotential in vacancy-rich ultrathin $CoSe_2$ nanosheets for water oxidation. *J Am Chem Soc* 2014; 136: 15670–15675.

29. Yi M, Shen Z. A review on mechanical exfoliation for the scalable production of graphene. *J Mater Chem A* 2015; 3: 11700–11715.

30. Niu L, Coleman JN, Zhang H, Shin H, Chhowalla M, Zheng Z. Production of Two-Dimensional Nanomaterials *via* Liquid-Based Direct Exfoliation. *Small* 2016; 12: 272–293.

31. Paton KR, Varrla E, Backes C, Smith RJ, Khan U, O'Neill A *et al*. Scalable production of large quantities of defect-free few-layer graphene by shear exfoliation in liquids. *Nat Mater* 2014; 13: 624–630.

32. O'Regan B, Grätzel M. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO_2 films. *Nature* 1991; 353: 737–740.

33. Manser JS, Christians JA, Kamat P V. Intriguing Optoelectronic Properties of Metal Halide Perovskites. *Chem Rev* 2016; 116: 12956–13008.

34. Liu J, Liu Y, Liu N, Han Y, Zhang X, Huang H *et al*. Metal-free efficient photocatalyst for stable visible water splitting *via* a two-electron pathway. *Science*

2015; 347: 970–4.

35. Castro Neto AH, Guinea F, Peres NMR, Novoselov KS, Geim AK. The electronic properties of graphene. *Rev Mod Phys* 2009; 81: 109–162.

36. Yeh T-F, Teng C-Y, Chen S-J, Teng H. Nitrogen-Doped Graphene Oxide Quantum Dots as Photocatalysts for Overall Water-Splitting under Visible Light Illumination. *Adv Mater* 2014; 26: 3297–3303.