Institution: University of St Andrews

Unit of Assessment: 5 - Biological Sciences

Title of case study: The 2A Protein Co-expression Technology for Biomedicine and Biotechnology

1. Summary of the impact

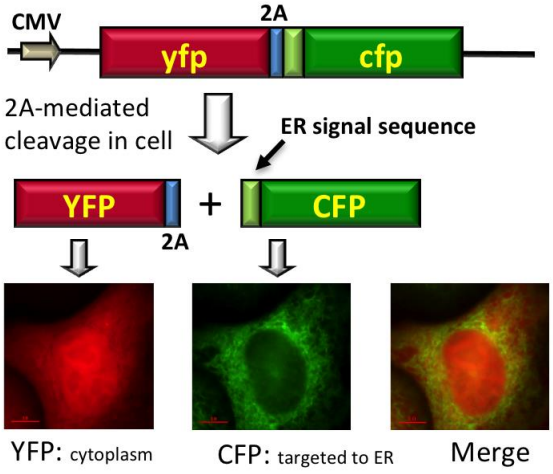
Co-expression of multiple proteins within the same cell is critical for success in many areas of biomedicine and biotechnology. This can now be readily accomplished by using 2A co-expression technology, developed by the Ryan Laboratory in St Andrews University. This technology has been critical in strategies for human gene therapies targeting cancer, production of induced human pluripotent stem cells for regenerative medicine, creation of transgenic animals and plants with improved nutritional properties and the production of high-value proteins for the pharmaceutical industry. Over 400 patent applications in the REF period utilise 2A, and multiple companies market products based on the technology.

2. Underpinning research

The 2A peptide sequence was first identified in picornaviruses such as Foot and Mouth Disease Virus (FMDV) and was known to be a site for peptide bond cleavage, but the mechanism was not understood. In 1994, a team led by **Prof Martin Ryan** (Professor in St Andrews from 1994 to present) demonstrated that FMDV 2A mediated a co-translational 'cleavage' (in the absence of other FMDV proteins) by creating an artificial polyprotein system. This comprised 2 genes fused together with, and without, an in-frame insertion of FMDV 2A between the two sequences, maintaining the single open reading frame. 'Cleavage' activity was determined by using (cell-free) translation systems *in vitro* and quantification of the translation products **[1]**. This was a key observation as it demonstrated that the 2A sequence could be used on its own to direct specific cleavage in a protein sequence. Using both *in vitro* and cell transfection data, the Ryan lab demonstrated in 1999 that this system could be used to co-express multiple, different, proteins in plants **[2]**. Furthermore, in the same year it showed that the 2 chains of the heterodimeric therapeutic cytokine interleukin-12 could be expressed efficiently by linking the 2 ORFs, *via* a 2A sequence, into a single gene **[3]**.

By this stage the Ryan lab had developed a model of 'cleavage: in fact due to a translational 'recoding' – which we termed 'ribosome skipping' in a very highly cited paper in 2001 **[4]**. This predicted that in the case of more complex artificial polyproteins, comprising multiple 2As, each 2A would function independently with no polarity effect with regards stoichiometry of the individual translation products.

This model of the 2A mechanism further predicted cleavage would occur within the ribosome: that any protein downstream of 2A would emerge from the ribosome with a nascent N-terminus. Therefore, if one included a co-translational signal sequence immediately downstream of 2A, it would be recognised by signal recognition particle (SRP) as if it were generated during 'normal' translation. The Ryan group showed in 2004 that this was the case: polyproteins were constructed comprising internal signal sequences such proteins could be both co-expressed and independently targeted to different sub-cellular sites (or secreted), greatly increasing the utility of the 2A co-expression system **[5]**.



The 2A peptide sequence allows efficient co-expression and differential targeting of multiple proteins

Subsequently in 2007-8, the Ryan group characterised a number of highly efficient 2A-like sequences from other viruses - which have been used to generate vectors for co-expression **[6]**, thus providing the technology to construct complex artificial 'self-processing' polyproteins that could encode the constituents of complete biochemical pathways, hetero-multimeric complexes etc. This technology has had huge impact in applications ranging from stem cell production

through transgenesis to the production of high value recombinant proteins.

3. References to the research

St Andrews Authors in bold; employment dates in St Andrews: Bruno 2003-07; De Felipe 2001-08; Donnelly 1994-98; El Amrani 1996-1998; Flint 1995-1996; Gani 1990-1998; Hughes 1994-98; Li 1998-2011; Lukashev 2006-2006; Luke 1994-present; Mehrotra 1994-1998; Ryan 1994-present.

These are all published in international, peer-reviewed journals. Total citations >450.

[1] Ryan, M.D. & Drew, J. (1994). Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO J*. 134, 928-933. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC394894/ (140 citations).

[2] Halpin, C., Cooke, S.E., Barakate, A., **El Amrani, A. & Ryan, M.D.** (1999). Self-processing polyproteins - a system for co-ordinate expression of multiple proteins in transgenic plants. *Plant J*. 17, 453-459. DOI: 10.1046/j.1365-313X.1999.00394.x (56 citations).

[3] Chaplin, P.J., Camon, E.B., Villarreal-Ramos, B., **Flint, M., Ryan, M.D.** & Collins, R.A. (1999). Production of interlukin-12 as a self-processing polypeptide. *J. Interferon & Cytokine Res*. 19, 235-241. DOI: 10.1089/107999099314162 (33 citations)

[4] Donnelly, M.L.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D. & Ryan, M.D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J. Gen. Virol*. 82, 1013-1025. http://vir.sgmjournals.org/content/82/5/1013.long (172 citations).

[5] El Amrani, A., Barakate, A., Askari, B.M., Li, X., Roberts, A.G., **Ryan, M.D.** & Halpin, C. (2004). Co-ordinate expression and independent subcellular targeting of multiple proteins from a single transgene. *Plant Physiol*. 135, 16-24. DOI: 10.1104/pp.103.032649 (31 citations).

[6] Luke, G.A., de Felipe, P., Lukashev, A., Kallioinen, S.E., **Bruno, E.A & Ryan, M.D.** (2008). The occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes. *J. Gen. Virol.* 89, 1036-1042. DOI: 10.1099/vir.0.83428-0 (22 citations).

4. Details of the impact

The 2A technology discovered and developed by Prof Martin Ryan in St Andrews allows the simultaneous expression of multiple proteins across a wide array of plant and animal cells and organisms. This greatly improves efficiency, simplifies the process of transgenesis, improves genetic stability of transgenics, expanding the types of genetic manipulation one can perform.

Specific impacts in the period 2008-2013 include:

- Development of a more nutritious variety of Golden Rice to treat vitamin A deficiency
- New cancer therapies based on co-expression of T cell receptor chains
- Efficient and predictable generation of induced Pluripotent stem cells by co-expression of transcription factors
- Numerous patented new techniques in healthcare, protein expression and plant biotechnology

The 2A system is superior to all other methods for the co-expression of multiple proteins, and facilitates much more complex strategies of transgenesis. In the words of the Director of the Tumor Immunology Group, Jonsson Comprehensive Cancer Centre, UCLA, USA in 2013 :

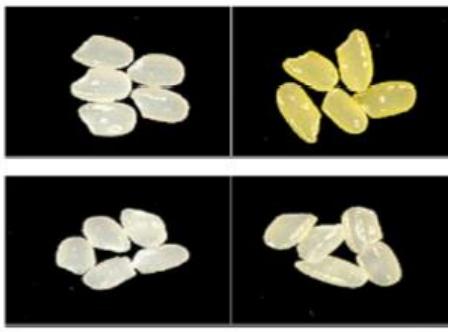
> "*The use of 2A sequences has allowed the planning and conduct of multiple gene transfer-based clinical trials. Without their description by Professor Ryan it is hard to conceive that we would be able to conduct clinical trials where several transgenes are expressed simultaneously and at high levels."* **[S1].**

Impact has been delivered in the following areas:

a) Plant Biotechnology to improve human nutrition.

Vitamin A deficiency (VAD) affects 100s of millions of people in Asia and Africa, and is implicated in ~2 million deaths per annum in children under 5. Some 500,000 children go blind each year due to VAD. Rice does not produce vitamin A, nor its precursor pro-vitamin A (β-carotene). The biosynthetic pathway leading to the synthesis of provitamin A was engineered into rice using 2A technology to make 'Golden Rice' with high levels of vitamin A by Korean scientists at the

National Institute of Agricultural Biotechnology. The 2A strategy proved much more successful than the use of an alternative method of co-expression, with at least 7 times higher levels of vitamin A (Ha et al. 2010, Plant Biotech. J. 8, 928-38). Patent protection was applied for in 2008 **[S4]**. Introduction of this strain of rice will help alleviate the severe health toll of VAD.



Golden rice made with 2A technology (top right) has more vitamin A than control rice (left) or golden rice made with alternative technologies (bottom right).

b) Cancer Therapy

New types of cancer treatment *via* gene therapy are revolutionising the treatment of the disease. Here, 2A is used to co-express the T-cell receptor TCRα and TCRβ chains (in transformed patient T-cells: *ex vivo* gene therapy) targeted against cancer antigens. Use of the 2A sequence ensures that both subunits of the receptor can be co-expressed at similar levels, minimising side effects. This strategy is now being used in the treatment of numerous types of cancer, notably metastatic melanoma. The Director of the Tumor Immunology Group, Jonsson Comprehensive Cancer Centre, UCLA, USA stated in 2013:

> "*This is particularly important for clinical trials based on the genetic engineering of the immune system to fight cancer. Many groups around the world are using 2A sequences for the expression of T cell receptors (TCR) for adoptive cell transfer immunotherapy in patients with cancer. This is because the two chains of the TCR need to be expressed efficiently and at similar levels. Furthermore, the use of 2A sequences has become such a routine that we are now planning the expression of a string of transgenes linked by 2A sequences in such clinical trials."* **[S1]**.

c) Regenerative Medicine

A second major impact in human health has been in the rapidly expanding field of regenerative medicine and transplantation. 2A has played a pivotal role in the co-expression of the four transcription factors required to produce induced pluripotent stem cells (iPSCs): one can now make patient-specific stem cells relatively simply to avoid problems of organ rejection or immune response. The Nobel Prize for Medicine 2012 was awarded to Prof Shinya Yamanaka of Kyoto University who is using the 2A system extensively in his work producing human iPSCs. He is establishing a bank of such cells to cover 90% of the Japanese population, thus providing near-universal tissue matches for medical purposes. The methodology utilizes 2A technology to streamline the process (eg Okita et al. (2011) *A more efficient method to generate integration-free human iPS cells*. Nature Methods 8, 409-412. doi:10.1038/nmeth.1591) and has been adopted by the wider scientific and commercial biomedical communities. The field is moving ahead very quickly because 2A peptide-linked cell reprogramming is *"Much more homogeneous compared with conventional systems"* according to an expert in the field from the University of Edinburgh **[S2]**. Indeed, the technology has been rolled out as a commercial product by companies such as ABM (www.abmgood.com), which markets "Yamanaka Factor Polycistronic Adenovirus" for iPCS generation that allows co-expression of the four transcription factors "*linked by protein 2A sequence which allows similar levels of transgene expression of four factors in target cells, thus streamlining iPSC reprogramming"* **[S5]**.

Figure from Cell Biolabs showing how the 4 transcription factors needed for iPSC reprogramming are linked by 2A sequences from different viruses

The company **iPierian** (www.ipierian.com), founded in 2008, is using 2A-based iPSC technology to develop new patient-specific cell lines to discover and develop new therapies for the treatment of neurodegenerative diseases such as Alzheimer's and Parkinson's. In 2009-11 it raised $78 million in start-up financing **[S6].** iPierian has entered a global licensing agreement with Kyoto University and Prof Yamanaka has joined iPierian's scientific board **[S7]**.

A wide variety of companies are now selling products for iPSC reprogramming that use 2A technology, including **Cell Biolabs** (www.cellbiolabs.com), **Biosettia,** (www.biosettia.com),

Stemgent (www.stemgent.com), **GenTarget,** (www.gentarget.com), **SBI System Biosciences** (www.systembio.com).

Together, these developments highlight the impact that facile multi-protein co-expression with 2A technology has had in the field of regenerative medicine.

d) Patents using 2A-based technology

Over 490 patent applications that use the 2A peptide technology have been made since Jan 2008. Over 100 patents using 2A technology have been granted in the same period, with applications aimed at cancer therapy, vaccine development, animal transgenesis, tumour suppression, novel antibiotics, targeted gene knockouts and protein production **[S8]**.

Conclusions

2A co-expression technology has therefore produced a highly substantial *direct* impact in many areas of modern molecular medicine / biotechnology since 2008. The underpinning research in St Andrews which demonstrated that 2A sequences were efficient self-cleaving entities was crucial to this impact. A 2012 report from the BBSRC, which funded much of the research in St Andrews, noted that:

> "*success has been notable in the fields of cancer treatment and regenerative medicine. 2A has been critical for the development of treatments for metastatic melanoma, colorectal cancer, synovial cell cancer and is entering trials for renal cancer. The sequence is also being used in the production of transgenic plants and tailor-made pluripotent stem cells*." **[S9]**

Finally, to quote the Vice Chair of Immunology, St Jude Children's Research Hospital, USA:

> "*It is rare for me to go to a meeting without at least one person approaching me and commenting in their successful use of the 2A system. I also know that several groups, including (Nobel Laureate) David Baltimore's, are using 2A-based vectors in pre-clinical studies. Thus, I have no doubt that* ***your 2A system has had a substantial impact*** *and that this will continue to grow over the next decade."* **[S3].**

5. Sources to corroborate the impact

[S1] E-Mail from the Director, JCCC Tumor Immunology Program Area – Jonsson Comprehensive Cancer Centre, UCLA, USA. Corroborates vital importance of 2A technology in situations where two or more proteins must be co-expressed for therapeutic goals.

[S2] Email from an independent expert from the University of Edinburgh. Corroborates vital importance of 2A technology for production of induced pluripotent stem cells.

[S3] E-mail from the Vice Chair of Immunology, St. Jude Children's Research Hospital, TN., USA. Corroborates impact of 2A technology on a wide range of new therapies for human disease.

[S4] Patent WO2009028903 A1, filed 2008. Corroborates the impact of 2A technology on production of new improved "Golden Rice" strains now in commercial development with large potential for improvement of human health.

[S5] Corroborates impact of 2A technology on products on sale for protein iPSC production. ABM product catalog http://www.abmgood.com/StemCell/dispOne-StemCell.php?page=MSDS&cart=0&csnn=30&csn=30&ssn=11344&dsn=11719&catno=000780A

[S6] Corroborates the size and strategy of iPierian, using 2A technology to generate iPSCs. http://www.xconomy.com/san-francisco/2011/05/13/ipierian-stem-cell-startup-with-big-science-big-bucks-axes-group-of-top-executives/

[S7] Corroborates the link between iPierian and Prof Yamanaka, who developed the use of 2A technology in iPSC production. http://www.fiercebiotech.com/press-releases/kyoto-university-ips-academia-japan-and-ipierian-announce-global-licensing-

[S8] Google patent search using term "2A peptide" and year range 2008-13. Corroborates impact of 2A technology on a wide range of applications in healthcare and protein production.

[S9] Corroborates the widespread application of 2A technology in applied fields. http://www.bbsrc.ac.uk/news/industrial-biotechnology/2012/120516-f-from-humble-beginnings.aspx