

## 2A - The “go-to” Technology for Transgene Co-expression

Ekaterina Minskaia<sup>1\*</sup> and Garry A. Luke<sup>2</sup>

<sup>1</sup>Institute of Molecular Medicine, Av. Prof Egas Moniz, 1649-028, Lisbon, Portugal

<sup>2</sup>Biomedical Sciences Research Complex, North Haugh, University of St. Andrews, St. Andrews, Fife, Scotland

### Abstract

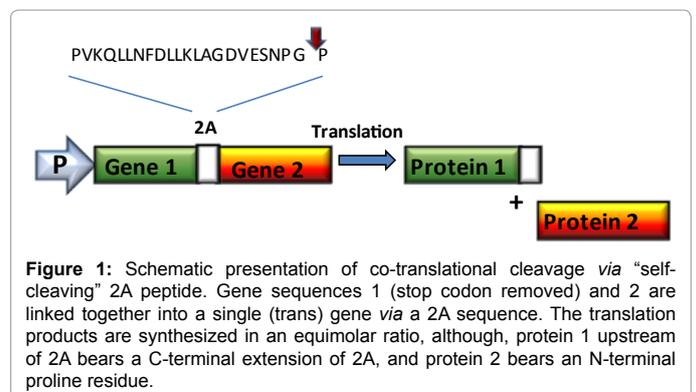
In order to co-express multiple genes for biotechnological and biomedical applications, several approaches have been used with varying degrees of success. Currently, internal ribosome entry site (IRES) elements and “self-cleaving” 2A peptides are the most widely used. The length of the IRES can be prohibitive and IRES-dependent translation of the second open reading frame is often significantly reduced. 2A peptides have gained in popularity due to their small size and ability to consistently produce discrete proteins at an equal level. Here, we promote the use of these sequences as the “go-to” technology for co-expression of multiple proteins.

**Keywords:** Protein co-expression; 2A; Biotechnology; Biomedicine

### Commentary

Many biotechnological and biomedical applications rely on the effective co-expression of multiple proteins. So far, multiple genes have been expressed via: (i) monocistronic vectors e.g. viral co-infection or co-transfection of plasmids expressing one protein each; fusion proteins; fusion proteins incorporating proteinase cleavage sites and (ii) polycistronic vectors where multiple genes are assembled either under the control of multiple promoters or a single promoter. IRES sequences have been used as a method to separate two coding sequences under the control of a single promoter. Despite their widespread use, they are relatively large (~600 base pairs) and expression of the downstream gene is significantly less efficient than the upstream gene [1]. A different approach using the 2A oligopeptide sequence allows multiple discrete proteins to be synthesized from a single strand of RNA, which also functions as a messenger RNA (mRNA). Not only is the 2A sequence smaller (54-174bp) than IRES elements, co-expression of proteins linked via 2A is independent of the cell type (cleavage activity is only dependent on eukaryotic ribosomes) and proteins are produced with equimolar stoichiometry [2-4]. The potential of this system is vividly demonstrated in the yeast *Pichia pastoris* by the functional co-expression of the carotenoid and violacein biosynthetic pathways encoded by nine genes from a polycistronic construct based on 2A peptides: the highest number of genes expressed in such a co-ordinated fashion so far [5]. IRES and 2A sequences have been used successfully in an impressive array of studies: at least 200 in the case of IRES elements and almost 900 in the case of 2A peptides making 2A the “go-to” option for protein co-expression [6-9]. Here we provide a short update on the history of 2A and cover some applications of 2A co-expression technology.

The Foot-and-mouth disease virus (FMDV) 2A sequence (hereafter “F2A”) mediates “self-processing” by a novel translational effect variously referred to as ‘ribosome skipping’, ‘stop-go’ and ‘stop-carry on’ translation [10-12]. 2A peptide cleavage has been studied in various cell types using various recombinant polyproteins and artificial reporter polyprotein systems comprising chloramphenicol acetyltransferase (CAT),  $\beta$ -glucuronidase (GUS), and fluorescent proteins (FPs e.g. GFP, RFP, YFP) [13-16]. It was demonstrated that F2A plus the N-terminal proline of the 2B downstream protein co-translationally ‘self-cleaved’ at the glycyl-prolyl pair site corresponding to the 2A/2B junction (LLNFDLLKLAGDVESNPG<sup>1</sup>P-) (Figure 1) [17-25]. Later work shows the length of the F2A used is also important for cleavage *in vitro* and *in vivo* - higher cleavage efficiency was reported when longer versions of 2A with N-terminal extensions derived from FMDV capsid protein 1D upstream of 2A were used [18,20,26-33]. Importantly, the co-



**Figure 1:** Schematic presentation of co-translational cleavage via “self-cleaving” 2A peptide. Gene sequences 1 (stop codon removed) and 2 are linked together into a single (trans) gene via a 2A sequence. The translation products are synthesized in an equimolar ratio, although, protein 1 upstream of 2A bears a C-terminal extension of 2A, and protein 2 bears an N-terminal proline residue.

translational nature of this cleavage means that by including various signal sequences within 2A polyproteins, either up- or downstream of 2A, proteins can be either targeted co-translationally to the exocytic pathway or post-translationally to different cellular compartments [34-36]. Of the many 2A peptides identified to date [10,27,37], four have been widely used in biotechnology and biomedicine: F2A, equine rhinitis A virus (“E2A”), porcine teschovirus-1 (“P2A”) and *Thosea asigna* virus (“T2A”) (Table 1) [2-4,36,38-43].

Finally, it should be noted that (i) 2A remains as a C-terminal extension of the upstream gene, and (ii) proline forms the N terminus of the downstream gene. The presence of N-terminal proline does not seem to affect proteins which are metabolically stable [44], but when the authentic C-terminus is required for activity or subcellular targeting of certain proteins, they should either be encoded at the C-terminus of the polyprotein or followed by cleavage sequences of the mammalian Kex2p homologue, furin (-<sup>1</sup>RRRR-, -<sup>1</sup>RKRR-, -<sup>1</sup>RRKR-), which removes the 2A “tag” [21,38]. The presence of 2A, however, can be used for detection of protein expression and localization using anti-2A antibodies [2,3].

**\*Corresponding author:** Ekaterina Minskaia, Institute of Molecular Medicine, Av. Prof Egas Moniz, 1649-028, Lisbon, Portugal, Tel: +351 217999566; E-mail: [minskayakat@hotmail.com](mailto:minskayakat@hotmail.com)

**Received** August 21, 2015; **Accepted** August 24, 2015; **Published** August 26, 2015

**Citation:** Minskaia E, Luke GA (2015) 2A – The “go-to” Technology for Transgene Co-expression. Single Cell Biol S1: 004. doi:10.4172/2168-9431.S1-004

**Copyright:** © 2015 Minskaia E, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviation	Source	2A/2A-like sequence
F2A	Foot-and-mouth disease virus (FMDV)	-PVKQLLNFDLLKLAGD <b>VESNPG P</b> -
E2A	Equine rhinitis A virus	-QCTNYALLKLAGD <b>VESNPG P</b> -
P2A	Porcine techovirus-1	-ATNFSLLKQAGD <b>VEENPG P</b> -
T2A	Thosea asigna virus	-EGRGSLTCGD <b>VESNPG P</b> -

**Table 1:** 2A and “2A-like” sequences used for protein co-expression. The – DxE<sub>N</sub>PGP- motif conserved among 2A/ 2A-like sequences is shown in red.

## Biotechnological and Biomedical Applications

F2A is the most widely used 2A sequence in plant biotechnology and has been used to target multiple proteins to various subcellular compartments [13,34,45,46], to improve disease resistance [47,48], drought-resistance [49] and nutritional value through metabolome engineering [10,50]. Vitamin A deficiency (VAD) is a major global health issue which affects hundreds of millions of people. This problem arises because rice, the staple food source in countries where VAD is prevalent, does not produce vitamin A or its precursor  $\beta$ -carotene, which have a number of vital functions in the body including growth. Since 2000, researchers have been engineering a transgenic variety of rice referred to as “golden rice” (*Oryza sativa*, GR) that includes the biosynthetic pathway for production of  $\beta$ -carotene [51,52]. Engineering the pathway into (carotenoid-free) rice endosperm requires two carotenoid biosynthetic genes, phytoene synthase (*psy*) and carotene desaturase (*crtl*) [53]. To avoid the problems of promoter interference (GR1 and 2 require two promoters), *psy* from Capsicum and *crtl* from *Pantoea*, were linked via 2A (*psy-F2A-crtl*; pPAC construct) or IRES (*psy-IRES-crtl*; pPIC construct) and placed under the control of the rice globulin promoter [50]. GR3 of transgenic PAC had a much more intense golden colour than did the PIC transformants, demonstrating that the 2A construct performed better than the IRES construct in terms of carotenoid production.

F2A and ‘2A-like’ sequences have been used extensively in genetic engineering of T cells for adoptive cell therapies [39-41], human stem cells [42,54,55] and in the induction of pluripotent stem cells [23,25,43]. For example, the genes encoding four transmembrane proteins necessary for the assembly of the CD3 complex were co-expressed from a polycistronic vector containing three 2As [36] and co-expression of four transcription factors in a [KLF4-E2A-OCT3/4-T2A-SOX2-P2A-c-Myc]-IRES-hrGFP construct resulted in the generation of induced pluripotent stem (iPS) cells from somatic cells [56]. High levels of functional monoclonal antibodies were produced by linking the antibody heavy and light chain sequences with F2A in the context of AAV-mediated gene transfer [21]. 2As were used to co-express tumor associated antigens (TCR)  $\alpha$ - and  $\beta$ -chains to treat metastatic melanoma, colorectal cancer, renal cell carcinomas and many other types of malignant diseases [57-59]. Significant anti-tumour responses were observed in the clinic using monoclonal antibodies by co-expressing cytotoxic T-lymphocyte-associated antigen (CTLA-4) heavy and light chains [60].

Finally, we think ‘2A-like’ sequences are able to function both as a signal sequence and as a translational recoding element - this leads to partitioning of the translation products between two sub-cellular sites (dual protein targeting). We have identified some 2A-like sequences at the N-terminus of NLRs in the genome of the purple sea urchin *Strongylocentrotus purpuratus* that were putative signal sequences. Constructs encoding wild-type [Sp2A-cherryFP-T2A-GFP] or a mutated, cleavage inactive form of Sp2A were used to transfect mammalian HeLa cells – with both constructs GFP was evenly distributed throughout the cell, wild-type Sp2A lead to cherryFP

localization throughout the cell and mutated Sp2A acting as a signal lead to cherryFP localization in the exocytic pathway [37,61].

In conclusion, with the number of studies that have successfully used F2A and ‘2A-like’ sequences approaching 1000, these small peptides are proving to be the ‘go-to’ technology for co-expression of multiple proteins.

## References

- Mizuguchi H, Xu Z, Ishii-Watabe A, Uchida E, Hayakawa T (2000) IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector Mol Ther 1: 376-82.
- de Felipe P, Hughes L, Ryan MD, Brown J (2003) Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. J Biol Chem 278: 11441-11448.
- Ryan MD, Drew J (1994) Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J 13: 928-933.
- Ryan MD, Donnelly M, Flint M, Cowton V, Luke G, et al. (2004) Foot-and-mouth disease virus proteinases. In: Foot and Mouth Disease Current Perspectives. Edited by Sobrino F, Domingo E. England: Horizon Bioscience: 53-76.
- Geier M, Fauland P, Vogl T, Glieder A (2015) Compact multi-enzyme pathways in *P. pastoris*. Chem Commun (Camb) 51: 1643-1646.
- Luke GA, Ryan M (2013) The protein co-expression problem in biotechnology and biomedicine: virus 2A and 2A-like sequences provide a solution. Future Virol 8: 983-996.
- Kerrigan J, Xie Q, Ames R, Lu Q (2011) Production of protein complexes via co-expression. Protein Expr and Purif 75: 1-14.
- Mokrejs M, Vopalensky V, Kolenaty O, Masek T, Feketova Z, et al. (2006) IRESite: the database of experimentally verified IRES structures. Nucleic Acids Res 34: 125-130.
- www.st-andrews.ac.uk/ryanlab/page10.htm
- Luke GA, Roulston C, Tilsner J, Ryan MD (2015) Growing Uses of 2A in Plant Biotechnology. In: Ekinci D (ed) Biotechnology. Rijeka: InTech Croatia: 165-193.
- Ryan MD, Donnelly M, Lewis A, Mehrotra A, Wilkie J, et al. (1999) A model for non-stoichiometric, co-translational protein scission in eukaryotic ribosomes. Bioorganic Chem 27: 55-79.
- Atkins J, Wills N, Loughran G, Wu C, Parsawar K, et al. (2007) A case for ‘StopGo’: Reprogramming translation to augment codon meaning of GGN by promoting unconventional termination (Stop) after addition of glycine and then allowing continued translation (Go). RNA 13: 1-8.
- Samalova M, Fricker M, Moore I (2006) Ratiometric fluorescence-imaging assays of plant membrane traffic using polyproteins. Traffic 7: 1701-1723.
- Provost E, Rhee J, Leach SD (2007) Viral 2A peptides allow expression of multiple proteins from a single ORF in transgenic zebrafish embryos. Genesis 45: 625-629.
- Trichas G, Begbie J, Srinivas S (2008) Use of the viral 2A peptide for bicistronic expression in transgenic mice. BMC Biol 6: 40.
- Burén S, Ortega-Villasante C, Ötvös K (2012) Use of the foot-and-mouth disease virus 2A peptide co-expression system to study intracellular protein trafficking in *Arabidopsis*. Plos ONE 7: e51793
- de Felipe P, Martin V, Cortes M, Ryan M, Izquierdo M (1998) Use of the 2A sequence from foot-and-mouth disease virus in the generation of retroviral vectors for gene therapy. Gene Ther 6: 198-208.
- Donnelly M, Hughes L, Luke G, Mendoza H, ten Dam E, et al. (2001) The ‘cleavage’ activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring ‘2A-like’ sequences. J Gen Virol 82: 1027-1041.
- de Felipe P, Hughes L, Ryan MD, Brown J (2003) Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. J Biol Chem 278: 11441-11448.
- Furler S, Paterna J, Weibel M, Bueler H (2001) Recombinant AAV vectors containing the foot- and-mouth disease virus 2A sequence confer efficient bicistronic gene expression in cultured cells and rat substantia nigra neurons. Gene Ther 8: 864-873.

21. Fang J, Qian J, Yi S, Harding TC, Tu G, et al. (2005) Stable antibody expression at therapeutic levels using the 2A peptide. *Nat Biotechnol* 23: 584-590.
22. Funston G, Kallioinen S, de Felipe P, Ryan MD (2008) Expression of heterologous genes in oncolytic adenoviruses using picornaviral 2A sequences that trigger ribosome skipping. *J Gen Virol* 89: 389-396.
23. Yang S, Cohen C, Peng P, Zhao Y, Cassard L, et al. (2008) Development of optimal bicistronic lentiviral vectors facilitates high-level TCR gene expression and robust tumor cell recognition. *Gene Ther* 15: 1411-1423.
24. Tan Y, Liang H, Chen A, Guo X (2010) Co-expression of double or triple copies of the rabies virus glycoprotein gene using a 'self-cleaving' 2A peptide-based replication-defective human adenovirus serotype 5 vector. *Biologicals* 38: 586-593.
25. Takata Y, Kishine H, Sone T, Andoh T, Nozaki M, et al. (2011) Generation of iPS cells using a BacMam expression system. *Cell structure and function* 36: 209-222.
26. Klump H, Schiedmeier B, Vogt B, Ryan M, Ostertag W, et al. (2001) Retroviral vector-mediated expression in HoxB4 in hematopoietic cells using a novel expression strategy. *Gene Ther* 8: 811-817.
27. Luke GA, de Felipe P, Lukashev A, Kallioinen S, Bruno E, et al. (2008) The occurrence, function, and evolutionary origins of '2A-like' sequences in virus genomes. *J Gen Virol* 89: 1036-1042.
28. Minskaia E, Nicholson J, Ryan M (2013) Optimisation of the foot-and-mouth disease virus 2A co-expression system for biomedical applications. *BMC Biotechnol* 13: 67.
29. Groot-Bramel-Verheije M, Rottier P, Meulenberg J (2000) Expression of a foreign epitope by porcine reproductive and respiratory syndrome virus. *Virology* 278: 380-389.
30. Kinsella T, Ohashi C, Harder A, Yam G, Li W, et al. (2002) Retrovirally delivered random cyclic peptide libraries yield inhibitors of interleukin-4 signalling in human B cells. *J Biol Chem* 277: 37512-37518.
31. Hahn H and Palmenberg A (2001) Deletion mapping of the encephalomyocarditis virus primary cleavage site. *J Virol* 75: 7215-7218.
32. Milsom M, Woolford L, Margison G, Humphries R, Fairbairn L (2004) Enhanced *in vivo* selection of bone marrow cells by retroviral-mediated co-expression of mutant O6-methylguanine-DNA-methyltransferase and HOXB4. *Mol Ther* 10: 862-873.
33. Lengler J, Holzmüller H, Salmons B, Gunzburg W, Renner M (2005) FMDV-2A sequence and protein arrangement contribute to functionality of CYP2B1-reporter fusion protein. *Anal Biochem* 343: 116-124.
34. El Amrani A, Barakate A, Askari B, Li X, Roberts A, et al. (2004) Coordinate expression and independent subcellular targeting of multiple proteins from a single transgene. *Plant Physiol* 135: 16-24.
35. Lorens J, Pearsall D, Swift S, Peelle B, Armstrong R, et al. (2004) Stable, stoichiometric delivery of diverse protein functions. *J Biochem Biophys Methods* 58: 101-110.
36. Szymczak A, Workman C, Wang Y, Vignali K, Dilioglou S, et al. (2004) Correction of multi-gene deficiency *in vivo* using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat Biotechnol* 22: 589-594.
37. Luke GA, Pathania U, Roulston C, De Felipe P, Ryan MD (2014) 'DxExNPGP - Motives for the motif'. In: *Recent Research Developments in Virology 9*, Research Signpost, India: 25-42.
38. Fang J, Yi S, Simmons A, Tu G, Nguyen M, et al. (2007) An antibody delivery system for regulated expression of therapeutic levels of monoclonal antibodies *in vivo*. *Mol Ther* 15: 1153-1159.
39. Szymczak A and Vignali D (2005) Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* 5: 627-638.
40. Huang X, Wilber A, Bao L, Tuong D, Tolar J, et al. (2006) Stable gene transfer and expression in human primary T cells by the Sleeping Beauty transposon system. *Blood* 107: 483-491.
41. Hart D, Xue S, Thomas S, Cesco-Gaspere M, Tranter A, et al. (2008) Retroviral transfer of a dominant TCR prevents surface expression of a large proportion of the endogenous TCR repertoire in human T cells. *Gene Ther* 15: 625-631.
42. Hsiao E, Yoshinaga Y, Nguyen T, Musone S, Kim J, et al. (2008) Marking embryonic stem cells with a 2A self-cleaving peptide: a NKX2-5 emerald GFP BAC reporter. *PLoS One* 3: 1-8.
43. Sommer C, Stadfeld M, Murphy G, Hochedlinger K, Kotton D (2009) Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* 27: 543-549.
44. Varshavsky (1992). The N-end rule. *Cell* 69: 725-735.
45. Halpin C, Cooke S, Barakate A, El Amrani A, Ryan M (1999) Self-processing 2A-polyproteins - a system for co-ordinate expression of multiple proteins in transgenic plants. *Plant J* 17: 453-459.
46. Samalova M, Fricker M, Moore I (2008) Quantitative and qualitative analysis of plant membrane traffic using fluorescent proteins. *Methods Cell Biol* 85: 353-380.
47. François I, Van Hemelrijck W, Aerts A, Wouters P, Proost P, et al. (2004) Processing in *Arabidopsis thaliana* of a heterologous polyprotein resulting in differential targeting of the individual plant defensins. *Plant Science* 166: 113-121.
48. Geu-Flores F, Olsen C, Halkier B (2009) Towards engineering glucosinolates into non-cruciferous plants. *Planta* 229: 261-270.
49. Kwon S, Hwang E, Kwon H (2004) Genetic engineering of drought-resistant potato plants by co-introduction of genes encoding trehalose-6-phosphate synthase and trehalose-6-phosphate of *Zygosaccharomyces rouxii*. *Korean J Genet* 26: 199-206.
50. Ha S, Liang Y, Jung H, Ahn M, Suh S, et al. (2010) Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. *Plant Biotechnol J* 8: 928-938.
51. Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, et al. (2000) Engineering the provitamin A ( $\beta$ -carotene) biosynthetic pathway into (Carotenoid-Free) rice endosperm. *Science* 287: 303-305.
52. Paine J, Shipton C, Chaggar S, Howells R, Kennedy M, et al. (2005) Improving the nutritional value of golden rice through increased pro-vitamin A content. *Nat Biotechnol* 23: 482-487.
53. Lu S and Li L (2008) Carotenoid metabolism: Biosynthesis, regulation, and beyond. *J Integrative Plant Biol* 50: 778-785.
54. Carey B, Markoulaki S, Hanna J, Saha K, Gao Q, et al. (2009) Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc Natl Acad Sci USA* 106: 157-162.
55. Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, et al. (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458: 771-775.
56. Shao L, Feng W, Sun Y, Bai H, Liu J, et al. (2009) Generation of iPS cells using defined factors linked via the self-cleaving 2A sequences in a single open reading frame. *Cell Res* 19: 296-310.
57. Johnson L, Morgan R, Dudley M, Cassard L, Yang J, et al. (2009) Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood* 114: 535-546.
58. Robbins P, Morgan R, Feldman S, Yang J, Sherry R, et al. (2011) Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. *J Clin Oncol* 29: 917-924.
59. Wang W, Hanada K, Feldman S, Zhao Y, Inozume T, et al. (2011) Development of a genetically-modified novel T-cell receptor for adoptive cell transfer against renal cell carcinoma. *J Immunol Methods* 366: 43-51.
60. Simmons A, Moskalenko M, Creson J, Fang J, Yi S, et al. (2008) Local secretion of anti-CTLA-4 enhances the therapeutic efficacy of a cancer immunotherapy with reduced evidence of systemic autoimmunity. *Cancer Immunol Immunother* 57: 1263-1270.
61. Roulston C, Luke G, de Felipe P, Ruan L, Cope J, et al. (2015) '2A'-like signal sequences mediating translational recoding: a novel form of dual protein targeting. *Traffic* (submitted).

This article was originally published in a special issue, **Progress in Single Cell Technology** handled by Editor(s), Dr. Yusuf Tutar, Cumhuriyet University, NJ, Turkey