Small proteins, big roles: The signaling protein Apela extends the complexity of developmental pathways in the early zebrafish embryo

*Michal Reichman-Fried and Erez Raz**

The identification of molecules controlling embryonic patterning and their functional analysis has revolutionized the fields of Developmental and Cell Biology. The use of new sequence information and modern bioinformatics tools has enriched the list of proteins that could potentially play a role in regulating cell behavior and function during early development. The recent application of efficient methods for gene knockout in zebrafish has accelerated the functional analysis of many proteins, some of which have been overlooked due to their small size. Two recent publications report on the identification of one such protein and its role in zebrafish embryogenesis. The protein, currently designated Apela, was shown to act as a secreted protein whose absence adversely affected various early developmental processes. Additional signaling proteins that have been identified in one of the studies are likely to open the way to unraveling hitherto unknown developmental pathways and have the potential to provide a more comprehensive understanding of known developmental processes.

Keywords:
Apela; Apelin; Aplnr; cardiogenesis; embryogenesis; gastrulation; GPCR; zebrafish

DOI 10.1002/bies.201400048

Institute of Cell Biology, Center for Molecular Biology of Inflammation, Münster, Germany

***Corresponding author:**
Erez Raz
E-mail: erez.raz@uni-muenster.de

Introduction

Embryonic development relies on a series of controlled events regulating cell proliferation, cell death, interactions among cells, cell differentiation, and migration of single and groups of cells. Together, these processes give rise to the three-dimensional shape of tissues and organs, as well as to the organization of the body plan. The identification of molecules controlling the execution of developmental programs progressed along two main paths. First, systematic random mutagenesis screens conducted in different model organisms led to the isolation of genes essential for normal embryonic development [1–5]. The functional analysis of the specific proteins and protein families identified in these screens yielded important clues concerning the mechanisms directing cell fate decisions and the regulation of cell behavior during embryogenesis. A second key approach for understanding the mechanisms controlling embryonic development was based on identification of genes with certain features of interest such as specific expression pattern, biochemical function, and/or interaction with other proteins; attractive candidates were chosen for targeted knock-out, knock-down, or over-expression experiments, followed by phenotypic characterization (e.g. for the functional analysis of the Goosecoid protein, [6–9]). The results of these two approaches revolutionized the fields of Developmental and Cell Biology and extended important implications for Medicine. Specifically, defects in pathways controlling developmental processes were found to constitute the basis for heritable medical conditions and diseases that develop during adult life. For example, the hedgehog pathway that controls segment polarity in *Drosophila* embryos [1, 10], was shown to be associated with the development of different malignancies in humans [11, 12].

The identification of molecules important for embryogenesis using these two methodologies – and especially the combination of the two – proved to be exceedingly

successful in providing entry points for studying a range of embryonic processes and the principles governing them. Despite the advances over recent decades in the detailed understanding of developmental processes, the complete spectrum of molecules playing a role in these events is not known. Multiple reasons contribute to those gaps in our knowledge; functional redundancy among molecules and pathways would mask potential phenotypes when not all redundant genes are co-targeted [13]. Similarly, inducing a phenotype by generating mutations in certain genes may prove difficult if those are present in multiple copies, making them practically refractory to identification by simple mutations (e.g. [14]). Genes that past screens have possibly missed include genes containing a very short open reading frame (ORF) [15]. Such genes could have escaped analysis because they were too small to be efficiently targeted in standard mutagenesis screens, and more difficult to predict computationally; small ORF genes could have thus been mistakenly considered as non-coding genes, which until recent years were regarded as of lower priority for targeted mutagenesis.

Ribosome-profiling and revised ORF annotations identifies novel zebrafish signaling molecules

To establish a more comprehensive list of pathways and molecules that function in animal development and physiology, Pauli et al. [16] carried out an unbiased search for novel translated ORFs in zebrafish. In their search, they employed ribosome profiling to predict the coding potential of transcripts that had previously been either not annotated or annotated as non-coding. Ribosome profiling allows the identification of ribosome-bound RNA fragments by deep sequencing as described earlier [17, 18]. This analysis led to the identification of hundreds of protein-coding genes, the majority of which exhibited sequence homology with proteins in other vertebrates. Significantly, 24 of the vertebrate-conserved proteins were shown to possess molecular features consistent with those of secreted signaling molecules, as they include a secretion leader protein sequence at their N-terminus and lack a trans-membrane domain. The list of small translated ORFs was further expanded as described in a work recently published by Bazzini et al. [19]. In this study, actively translated ORFs from zebrafish and human were identified by ribosome profiling. In addition, the use of a new computational tool facilitated the selection of an overlapping set of zebrafish and human peptides exhibiting conserved sequences. Sixty-three conserved short zebrafish peptides and 173 human peptides were identified in this way, among which 23 and 7, respectively, were identified by both methods. The future functional analysis of these new molecules will likely have a major impact on our understanding of gaps still existing in known developmental pathways on the one hand, uncovering previously unknown pathways that function during embryonic and later stages of development on the other.

Apela, a new signaling molecule that controls early embryonic development

As a proof-of-principle, Pauli et al. [16] analyzed one of the candidate proteins that they named Toddler (Fig. 1). This previously unknown protein was independently identified and named ELABELA in a parallel study [20]. The signaling protein was recently designated "Apela" (Apelin receptor early endogenous ligand) by the HUGO Gene Nomenclature Committee HGNC, and hereafter will be referred to by this name. The zebrafish *apela* gene – which was previously considered noncoding [21, 22] – was found to encode a 58-amino acid ORF that showed strong phylogenetic conservation with the chick, frog, mouse, and human counterparts, in particular within the C-terminal part of the secreted protein. In the other vertebrate species examined, Apela consists of 54 amino acids [20]. Remarkably, both studies described similar phenotypes induced in the absence of the encoded protein, and identify Apelin receptor as its cognate receptor. The authors introduced mutations in the gene employing zinc-finger (ZFN) [20] and transcription activator-like effector nuclease (TALEN) [16] mutagenesis, and found that most of the homozygous mutant animals died at the end of embryogenesis (5–7 days of development). The function of *apela* appeared to be required only during embryonic stages, because providing mutant embryos with the wild-type gene product during early embryogenesis allowed treated animals to reach adulthood.

Apela function is required during gastrulation and heart formation

A more detailed description of the protein function was unraveled as the phenotypes of the mutant embryos were examined. Pauli et al. [16] focused primarily on the effect that loss of Apela function had on the process of gastrulation, while Chng et al. [20] reported on the effect of the mutation on cardiovascular development. Both groups, however, noted similar ranges of defects at the early developmental stages. Analyzing morphology and early gene expression patterns during gastrulation, Pauli et al. [16] observed normal early cell specification, but defective cell movement. Specifically, delayed cell movement in the gastrulating embryos was observed during the spreading of the cells toward the vegetal pole in the process of epiboly [16], as well as during internalization of the mesendodermal cells and their migration in the direction of the animal pole. Interestingly, unlike some secreted attractive signals that guide migration and are thus expressed at the position of the migration target, *apela* transcripts are ubiquitously expressed during early development. Assuming a simple spatial and temporal correlation between transcription and the secretion of the protein, these findings argue against a role for Apela in guiding cell migration. Consistent with this notion, is the finding that both ubiquitous and localized expression of Apela effectively restored cell migration in *apela* mutants. Based on the phenotype they observed during zebrafish gastrulation, Pauli et al. concluded that Apela functions as a promoter of cell motility. While Chng et al. also detected defects in early

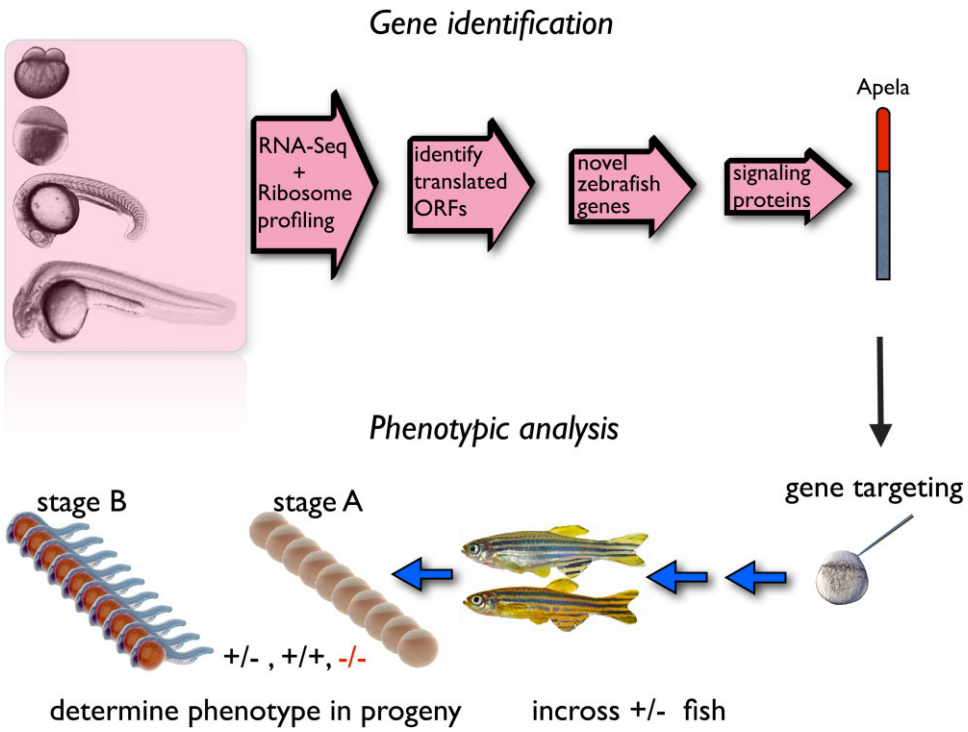

Figure 1. The workflow carried out by Pauli et al. [16] to identify and functionally analyze the Apela protein. Targeted mutagenesis of genes identified in molecular screens in zebrafish as performed for the *apela* gene using the ZFN [20] and TALEN [16] technology, is currently done primarily by employing the "clustered, regularly interspaced, short palindromic repeats" (CRISPR)-based method [39].

morphogenetic movements, they mainly reported on differentiation defects in the endodermal lineage, and on the failure of mutant embryos to form a normal heart, showing rudimentary heart, or no heart at all [20]. These findings highlight the pleiotropic nature of the phenotype and the fact that Apela is required in a range of processes occurring during embryonic development. The precise role of Apela in these processes could still constitute a similar function that is required in both motile and differentiating cells of the developing zebrafish embryos. A prominent feature of the *apela* mutation, as noticed in both studies, was the variability in the severity of the effect as manifested in the wide spectrum of heart anomalies, tail development, and trunk truncation. Further, the variable phenotypes ranged from early embryonic death to occasional survival of *apela* mutants to adulthood with apparently normal morphology.

Apela signals via the Apelin receptor (Aplnr)

Several lines of evidence led both groups to identify the G-protein coupled receptor (GPCR) for Apelin as the receptor for the ligand. Firstly, the phenotypes observed upon loss of function of Apela resemble that induced following removal of the Apelin receptor (of which the zebrafish genome harbors two paralogs, i.e. *aplnra* and *aplnrb*). In addition, the human Apela protein showed effective binding to cells engineered to express the Apelin receptors in an in vitro setting [20], and over-expression of the ligand in embryos led to internalization of the labeled receptor in vivo [16], an event that commonly marks the interaction between a ligand and its GPCR. The significance of identifying Aplnr as Apela's receptor lies in the large array of functions already described for this receptor. For example Apelin receptors have been implicated in central and peripheral responses to multiple homeostatic perturbations that include regulating cardiovascular function and angiogenesis [23, 24]. Whereas the two groups focused on two aspects of embryonic development, the wide-spread expression of the receptors and their ligand suggests that Apelin-receptor signaling may be required for additional events during embryogenesis. Such an additional function for the receptor has already been defined in the zebrafish lymphatic system [25, 26], and in diverse processes in other organisms [23]. It is likely therefore, that a large number of unexplored functions are yet to be revealed. Examining developmental events in *apela* mutant embryos should expand the range of processes known to be controlled by the receptor, providing important hints concerning the precise developmental role of the Apelin signaling pathway. Interestingly, unlike in the mouse [27], in zebrafish the two ligands of the Apelin receptor, Apela and Apelin, are expressed successively during early embryonic development, such that *apela* transcripts arise first and those of *apelin* are detected 5 hours later [16, 20, 28, 29]. Thus, studying the Apelin signaling pathway in zebrafish should enable the uncoupling of early from later processes that are regulated by this pathway. For example as Apelin knock out mice show a range of mild heart defects [30], it is likely that the presence

of Apela (referred to as "ende" in mice [27]) during those stages allows most of the heart formation process to unfold properly.

Conclusions and prospects

Employing the recent advances in reverse genetic techniques in zebrafish to analyze the function of the newly identified putative signaling molecules, novel developmental signals, and pathways will probably be uncovered. As demonstrated for the *apela* locus, phenotypic analysis of the new mutants can provide novel insights into the mechanisms controlling embryonic development. Bearing in mind the phenotypes induced by loss of the Apela protein, the next step in the analysis would be the thorough characterization of the specific effect on the development of different tissues and organs of the embryo. As the processes regulated by the receptor are identified and thoroughly described in zebrafish [28, 29], it would be important to search for a common thread that ties the actions of Apela protein in all of them. One could then determine whether the events governed by this signaling pathway primarily control motility (as suggested by Pauli et al. [16]), regulate cell differentiation (as suggested by Chng et al. [20]) or whether they reflect a cross talk between the cell movement and fate determination. Additional processes, such as cell proliferation, appear to be directly or indirectly affected in the absence of Apela as can be deduced from the reduction in endodermal cell number observed by both groups. Determining whether the requirement for Apela is cell autonomous will shed light on the function of the peptide. Apela function could be essential for establishing a permissive environment within which cells of different types would be able to carry out their development program properly. Alternatively, the intracellular signaling in response to Apela binding could be required for normal differentiation and/or migration in a cell-autonomous manner. This issue was previously investigated in the context of cardiac development, where it was found that Aplnr functions in a non-autonomous manner upon Apelin binding [31]. It is however formally possible that Apela and Apelin induced different signals upon engagement of Aplnr, leaving the question to be directly addressed for each process. In addition, while Chng et al. demonstrate the expression of Apela in human embryonic stem cells, the expression of Aplnr in these cells was not reported [20], giving rise to the idea that Apela might interact with more than one receptor. Together, the elucidation of the cell autonomy nature of the receptor-ligand interaction would allow the classification of the effects as primary versus secondary among the wide spectrum of Apela-induced phenotypes. It would then be important to integrate the action of the Apelin receptor-controlled biochemical cascade with the currently known mechanisms controlling the specific events in embryogenesis. For example, how does the signaling pathway downstream of Apela interact with pathways essential for general cell motility [32–34] and with pathways specifically associated with cell movement in gastrulation [35, 36]? Similarly, the link between Apela signaling pathway and developmental pathways controlling endoderm differentiation could be explored [37, 38]. Considering the wide range of *apela*-induced phenotypes in zebrafish and the broad array of Apelin receptor functions in mice [23], it is conceivable that Apela signaling controls common developmental and physiological events. While it is clear that Apela is essential during early zebrafish embryogenesis, the question of whether Apelin-receptor signaling plays a role in later developmental stages is still open. Analysis of an Apelin-receptor null mutant at adult stages would therefore be informative.

Acknowledgments

We acknowledge the support of the University of Münster, the Cells in Motion (CIM) excellence cluster, the German research foundation (DFG), and the European Research Council (ERC). We also thank Andrea Pauli and Bruno Reversade for their helpful comments on the manuscript.

References

1. **Nusslein-Volhard C, Wieschaus E**. 1980. Mutations affecting segment number and polarity in Drosophila. *Nature* **287**: 795–801.
2. **Haffter P, Granato M, Brand M, Mullins MC**, et al. 1996. The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**: 1–36.
3. **Driever W, Solnica-Krezel L, Schier AF, Neuhauss SC**, et al. 1996. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**: 37–46.
4. **Stanford WL, Cohn JB, Cordes SP**. 2001. Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* **2**: 756–68.
5. **Li MA, Pettitte SJ, Yusa K, Bradley A**., 2010. Genome-wide forward genetic screeens in mouse ES cells. In: Wassarman P, Soriano PM, ed.; *Guide to Techniques in Mouse Development, Part B: Mouse Molecular Genetics* San Diego, USA: Elsevier, p. 217–41.
6. **Cho KW, Blumberg B, Steinbeisser H, De Robertis EM**. 1991. Molecular nature of Spemann's organizer the role of the Xenopus homeobox gene goosecoid *Cell* **67**: 1111–20
7. **Rivera-Pérez JA, Mallo M, Gendron-Maguire M, Gridley T**, et al. 1995. Goosecoid is not an essential component of the mouse gastrula organizer but is required for craniofacial and rib development. *Development* **121**: 3005–12.
8. **Yamada G, Mansouri A, Torres M, Stuart ET**, et al. 1995. Targeted mutation of the murine goosecoid gene results in craniofacial defects and neonatal death. *Development* **121**: 2917–22.
9. **De Robertis EM, Larrain J, Oelgeschlager M, Wessely O**. 2000. The establishment of Spemann's organizer and patterning of the vertebrate embryo. *Nat Rev Genet* **1**: 171–81.
10. **Ingham PW, Nakano Y, Seger C**. 2011. Mechanisms and functions of Hedgehog signalling across the metazoa. *Nat Rev Genet* **12**: 393–406.
11. **Bailey JM, Singh PK, Hollingsworth MA**. 2007. Cancer metastasis facilitated by developmental pathways: Sonic hedgehog, Notch, and bone morphogenic proteins. *J Cell Biochem* **102**: 829–39.
12. **Xie J, Murone M, Luoh SM, Ryan A**, et al. 1998. Activating smoothened mutations in sporadic basal-cell carcinoma. *Nature* **391**: 90–2.
13. **Kafri R, Levy M, Pilpel Y**. 2006. The regulatory utilization of genetic redundancy through responsive backup circuits. *Proc Natl Acad Sci USA* **103**: 11653–8.
14. **Woollard A**. 2005. Gene duplications and genetic redundancy in *C. elegans*. *WormBook* **25**: 1–6.
15. **Slavoff SA, Mitchell AJ, Schwaid AG, Cabili MN**, et al. 2013. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol* **9**: 59–64.
16. **Pauli A, Norris ML, Valen E, Chew GL**, et al. 2014. Toddler: an embryonic signal that promotes cell movement via Apelin receptors. *Science* **343**: 1248636.
17. **Chew GL, Pauli A, Rinn JL, Regev A**, et al. 2013. Ribosome profiling reveals resemblance between long non-coding RNAs and 5′ leaders of coding RNAs. *Development* **140**: 2828–34.

18. **Ingolia NT, Lareau LF, Weissman JS**. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**: 789–802.
19. **Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD**, et al. 2014. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J* **33**: 981–93.
20. **Chng SC, Ho L, Tian J, Reversade B**. 2013. ELABELA: a hormone essential for heart development signals via the apelin receptor. *Dev Cell* **27**: 672–80.
21. **Ulitsky I, Shkumatava A, Jan CH, Sive H**, et al. 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**: 1537–50.
22. **Guttman M, Donaghey J, Carey BW, Garber M**, et al. 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**: 295–300.
23. **O'Carroll AM, Lolait SJ, Harris LE, Pope GR**. 2013. The apelin receptor APJ: journey from an orphan to a multifaceted regulator of homeostasis. *J Endocrinol* **219**: R13–35.
24. **Barnes G, Japp AG, Newby DE**. 2010. Translational promise of the apelin-APJ system. *Heart* **96**: 1011–6.
25. **Karpinich NO, Caron KM**. 2014. Apelin signaling: new G protein-coupled receptor pathway in lymphatic vascular development. *Arterioscler Thromb Vasc Biol* **34**: 239–41.
26. **Kim JD, Kang Y, Kim J, Papangeli I**, et al. 2014. Essential role of Apelin signaling during lymphatic development in zebrafish. *Arterioscler Thromb Vasc Biol* **34**: 338–45.
27. **Hassan AS, Hou J, Wei W, Hoodless PA**. 2010. Expression of two novel transcripts in the mouse definitive endoderm. *Gene Expr Patterns* **10**: 127–34.
28. **Zeng XX, Wilm TP, Sepich DS, Solnica-Krezel L**. 2007. Apelin and its receptor control heart field formation during zebrafish gastrulation. *Dev Cell* **12**: 391–402.
29. **Scott IC, Masri B, D'Amico LA, Jin SW**, et al. 2007. The g protein-coupled receptor agtrl1b regulates early development of myocardial progenitors. *Dev Cell* **12**: 403–13.
30. **Charo DN, Ho M, Fajardo G, Kawana M**, et al. 2009. Endogenous regulation of cardiovascular function by apelin-APJ. *Am J Physiol Heart Circ Physiol* **297**: H1904–3.
31. **Paskaradevan S, Scott IC**. 2012. The Aplnr GPCR regulates myocardial progenitor development via a novel cell-non-autonomous, Galpha(i/o) protein-independent pathway. *Biol Open* **1**: 275–85.
32. **Parent CA, Weiner OD**. 2013. The symphony of cell movement: how cells orchestrate diverse signals and forces to control migration. *Curr Opin Cell Biol* **25**: 523–5.
33. **Abu Shah E, Keren K**. 2013. Mechanical forces and feedbacks in cell motility. *Curr Opin Cell Biol* **25**: 550–7.
34. **Doyle AD, Petrie RJ, Kutys ML, Yamada KM**. 2013. Dimensions in cell migration. *Curr Opin Cell Biol* **25**: 642–9.
35. **Yin C, Ciruna B, Solnica-Krezel L**. 2009. Convergence and extension movements during vertebrate gastrulation. *Curr Top Dev Biol* **89**: 163–92.
36. **Rohde LA, Heisenberg CP**. 2007. Zebrafish gastrulation: cell movements, signals, and mechanisms. *Int Rev Cytol* **261**: 159–92.
37. **Ober EA, Field HA, Stainier DY**. 2003. From endoderm formation to liver and pancreas development in zebrafish. *Mech Dev* **120**: 5–18.
38. **Zorn AM, Wells JM**. 2007. Molecular basis of vertebrate endoderm development. *Int Rev Cytol* **259**: 49–111.
39. **Blackburn PR, Campbell JM, Clark KJ, Ekker SC**. 2013. The CRISPR system – keeping zebrafish gene targeting fresh. *Zebrafish* **10**: 116–8.