Long non-coding RNAs have become the focus of considerable interest over the past few years. Intriguing novel functions have been reported for lincRNAs. Three recent papers identify lincRNAs that work in a more conventional way—encoding protein—in each case a small polypeptide with an interesting biological activity (Magny et al, 2013; Pauli et al, 2014), (Bazzini et al, 2014).

These stories have a precedent, or two. In 2004, the Drosophila polar granule component gene (pgc) was reported to function as a non-coding RNA that acted in the embryonic primordial germ cells to prevent transcription of the zygotic genome (Martinho et al, 2004). pgc RNA localizes to the nascent germ cells in the embryo to transiently block activation of RNPoII. A few years later, in 2008, the Nakamura and Ladurner laboratories reported that the functional product of the pgc gene was a peptide that blocks RNPoII by preventing an activating phosphorylation event mediated by P-Telb (Hanyu-Nakamura et al, 2008; Timinszky et al, 2008). A second “former lincRNA” produced by the tarsal-less/polished rice/mille-pattes gene turns out to encode small peptides that control epithelial morphogenesis in Drosophila and Tribolium (Savard et al, 2006; Galindo et al, 2007; Kondo et al, 2007).

Intriguingly, this peptide promotes N-terminal processing of the transcription factor Shavenbaby, converting it from a repressor to an activator (Kondo et al, 2010). The recent report from the Couso laboratory built on their previous work on tarsal-less to search for additional Drosophila transcripts that might encode small peptides. They identified a lincRNA that is expressed in muscle and encodes small peptides (Magny et al, 2013). Interestingly, these peptides are related to the vertebrate peptides Sarcolipin and Phospholamban in sequence and predicted structure. Mutants lacking the fly lincRNA, which they name sarcolamban, show a defect in cardiac function. Based on the known role of the human peptides in calcium uptake by this family of small peptides, the Giraldez and Rajewsky laboratories set out to survey Zebrafish lincRNAs for protein-coding potential. Both groups made use of ribosome profiling, a method that allows high-resolution mapping of ribosome-bound RNA fragments by deep sequencing (Ingolia et al, 2009). The Schier laboratory specifically looked for novel secreted proteins (Pauli et al, 2014). They found 700 predicted open reading frames that had not been annotated previously as protein-coding transcripts. Over 80% were conserved in other vertebrates and many encode polypeptides of considerable length—so it is unclear why they were missed before. Some of these transcripts had been annotated as lincRNAs. 28 of these were predicted to be new secreted proteins, with signal peptides, but lacking transmembrane domains. The new paper focused on the role of one of these loci, now named toddler, which encodes a secreted peptide. Using TALEN technology to produce mutants disrupting the peptide coding sequence, Pauli et al (2014) provide evidence that the peptide functions as a signal to promote cell motility in the early fish embryo. Toddler peptide, also known as ELABELA, acts as an activator of a G protein-coupled receptor called Apelin to promote cell movements required for heart development (Chng et al, 2013; Pauli et al, 2014).

The Giraldez and Rajewsky laboratories used ribosome profiling and developed new computational methods to identify Zebrafish lincRNAs that might encode novel short proteins (Bazzini et al, 2014). The
ORFs have been assigned to only a few of these former lincRNAs, but the versatile tools available for genome manipulation should allow a rapid follow-through. We can expect to be hearing a lot about the small protein world in years to come.

Conflict of interest
The author declares that he has no conflict of interest.

References
Drosophila RNA polymerase II transcription by a Drosophila oligopeptide. PLoS ONE 3: e2506