

# An excitingly predictable 'omic future

Joanna Asprer\*

Developmental biology is an old field of study that finds its origins in the 18th century, when Wilhelm Roux ablated cells from two- or four-cell frog embryos and observed that the remaining cells could not give rise to an entire embryo (Roux, 1888). It is now more than a century after the publication of that seminal work, and we can attest to the saying that the more things change, the more they stay the same; for here we are, still preoccupied with the same issue of stem cell pluripotency.

That matter aside, developmental biology has definitely come a long way. Whereas before, Roux and his contemporaries sought to study embryonic development through microsurgical manipulation of early embryos, we now alter model organisms on a genetic level, creating knockouts and transgenics, and we routinely analyze them on a molecular level by performing quantitative reverse transcriptase polymerase chain reactions (qRT-PCRs), chromatin immunoprecipitation (ChIP) assays, co-immunoprecipitation (Co-IP), immunohistochemical analyses and so on.

The most recent additions to our arsenal, microarray and next-generation sequencing, have been so pivotal to the progress of the field that they have ushered in an entirely new era: the 'omics era. We are no longer limited to the analysis of one or several genes at a time; instead, we attempt to take in the full breadth of the changes that occur as development progresses, from the level of organisms, organs, tissues, cells and molecules, to the level of genomes, epigenomes and transcriptomes.

## We are no longer limited to the analysis of one or several genes at a time

Interestingly, the field of developmental biology was slow to incorporate those technologies. The pre-translational 'omics era began in 1995, when the laboratory of Patrick Brown published a study about 45 *Arabidopsis* genes that were analyzed by microarray (Schena et al., 1995). The technology was quickly applied by those in the fields of cancer genetics, medical genetics, immunology and even drug development. It was only in 1999 that the first developmental papers studying *Drosophila* genes by microarray were published (Bryant et al., 1999; White et al., 1999).

The lag seemed concordant with the general inaccessibility of the technology in its infancy. In fact, in the 1990s, scientists complained about restricted access to microarray technology, citing high cost, limited availability and patent disputes (Fox, 1999). Biotech companies that had been generating expensive arrays for industry labs responded by producing lower cost arrays for academic labs. Meanwhile, institutions began to build core facilities to make the technology available to more scientists. Today, a PubMed search using the keywords 'microarray and embryonic development' yields close to 2000 results, and this is by no means an exhaustive search of all developmental studies that have made use of array technology.

Next-generation sequencing technology has also existed in some form as far back as the 1990s, when it was only useful to developmental biologists for generating expressed sequence tag (EST) libraries. Soon after the technology was commercialized in 2005, it was used to catalog developmental microRNAs (miRNAs) in at least two model organisms (Chen et al., 2005; Mineno et al., 2006), an undertaking that has proven to be essential for the continued progress of developmental biology. However, medical genetics, cancer genetics, microbiology and virology made much bigger strides using the same technology. This may be explained by the fact that the technology was primarily designed to sequence genomes and is understandably valuable in fields in which there is great genetic variation within sample populations as well as a constant supply of new genomes to sequence.

For developmental biologists who make use of isogenic strains of organisms with fully sequenced genomes, it was the subsequent application of next-generation sequencing to the quantitative analysis of gene expression and gene regulation via RNA sequencing (RNA-seq) and ChIP sequencing (ChIP-seq) that triggered its complete assimilation into the field. Still, even with the development of more relevant applications of the technique, there are currently fewer than 100 papers on PubMed that can be located with the search words 'RNA seq' or 'ChIP seq' and 'embryonic development'. This is again not an exhaustive list, but it gives us an idea that next-generation sequencing is not as widespread or as accessible as we would like it to be.

Both next-generation sequencing and microarrays are definitely not seen as routine tools on the level of even qRT-PCRs, at least not in the typical lab. Often, they are used as screens, performed with great cost and effort once in a blue moon. The intervening time is used to parse through the data and validate a multitude of individual results using more traditional techniques. This trend will continue for some time yet, but we are now at another turning point; technological advancements with semiconductor chips allow us to sequence a human genome in one day for a mere \$1000: a far

### 'Developments in development' essay competition

The essay 'An excitingly predictable 'omic future' was the winning entry in an essay competition on the Node, the community site for developmental biologists. Entrants were asked to write about 'Developments in development', and submissions discussed a variety of themes – such as epigenetics, data analysis, model organisms, or stem cells – all related to the future of developmental biology.

After a round of judging, two essays were nominated, and a final voting round on the Node determined the winning essay, written by Joanna Asprer. Joanna has a doctorate in developmental biology from Baylor College of Medicine and has worked on brain, spinal cord and inner ear development. She is currently a postdoc in Uma Lakshmi's lab at Life Technologies where she studies embryonic and induced pluripotent stem cells.

The runner-up in the essay competition was Máté Varga, with an essay about genetic engineering. For more information on the competition, and to see both nominated essays, visit the Node: <http://thenode.biologists.com/essay-competition-nominees/>.

cry from the decade and billions of dollars that were needed to complete the Human Genome Project in the year 2000.

With the growing accessibility of next-generation sequencing, more and more labs will do it more and more often. At some point, the affordability of new technologies will take high-throughput experiments from the domain of core facilities and give direct, in-lab access to a wider variety of scientists. Someday, hopefully soon, we will be able to do the high-throughput experiments in our own labs, comparing what makes Cell A different from Cell B, answering questions about how Process C changes the behavior or identity of Cell A, and finding out what DNA or RNA sequences Protein D binds to in order to induce the development of Cell B from Cell A in response to Process C.

Based on what happened in response to the introduction of microarray cores, we can predict that this unparalleled access to high-throughput technologies will accelerate progress in the field. True, the democratization of high-throughput experiments is bound to bring with it a host of problems associated with inexperience, such as misuse of equipment, faulty experimental design or misinterpretation of results. However, with time it will encourage a wider distribution of expertise which, combined with each lab having greater control over the equipment, will allow the execution of more experiments with possibly more complex experimental designs that will allow us to answer questions we could not tackle before.

### ...in my opinion, the future of developmental biology lies in the inevitable integration of the great trends that saw their beginnings in the last two decades...

So much data will be generated in this sequencing revolution that our search for the needle in a haystack will evolve into a search for the crucial pieces of information hidden in terabytes of data. As the challenge shifts from data-gathering to data-mining, bioinformatics will be more important than ever before. There will be great demand for recruitment of bioinformatics experts in order to address developmental issues, and even classical developmental biologists will be compelled to get training, if only to be able to design experiments properly and understand their collaborators sufficiently.

With teams of developmental biologists and bioinformatics experts working together, I envisage a scenario in which we would be able to speed up the process of mapping mutations generated by random mutagenesis using whole-genome sequencing. It may be a long while yet before this becomes a viable option in the fly field, but it might already represent an economical alternative to maintaining large mouse colonies for long periods of time while doing segregation analysis.

Along those same lines, many mouse labs have experienced the tragedy of running across a mutant with an exciting phenotype only to observe that the phenotype has low penetrance, or, worse, has disappeared with a transition from one genetic background to another. In this situation, it is quite likely that there is a modifier influencing the gene of interest. At the moment, we are given only two extreme options: throw in the towel and move on to something more tractable, or bull-headedly and blindly backcross into a defined background or two until the phenotype resurfaces or exhibits itself consistently. Soon, we may have the third, faster and more productive option of sequencing genomes and finding the genetic change or polymorphism that segregates with the phenotype or the lack thereof. Thus, the lab tragedy can turn into a serendipitous discovery about the pathway the lab is working on.

Clearly, affordable next-generation sequencing will be a great boon to basic scientists, but the 'omics era comes at a time when we are experiencing a strong push for translational or disease-relevant science, courtesy of the grim economic outlook and the corresponding decrease in research funding and increase in demand for responsible use of public funds. The alignment of technological availability and the renewed thrust towards improving human health will work out very well for the field of medical genetics, in which the recent launch of whole-exome sequencing for diagnostic purposes could reveal novel disease-causing mutations. This development promises to benefit other fields, including ours, as many of the discovered mutations will inevitably affect developmental genes and the human phenotype-genotype correlations will provide a nice complement to the studies carried out with model organisms.

Meanwhile, in developmental cell biology, great emphasis will continue to be placed on the study of different kinds of stem cells and their differentiation into cells that can be used for therapeutic purposes. This endeavor can benefit from high-throughput analyses as well: at present, we rely on one or several markers to characterize pluripotent stem cells or the differentiated cells that we have turned them into, but we know that expression of one gene does not a cell type make. Even our so-called markers tend to be expressed by multiple cell types, all of which represent possible end points of pluripotent stem cell differentiation. Moreover, expression of a certain marker does not guarantee that an induced cell will behave and function exactly like the predicted cell type. Thus, there will be a point at which it will be unacceptable to use only a few markers in identifying cells for therapeutic use and, instead, transcriptome fingerprinting will become the norm.

I have no doubt that many completely new things will eventually revolutionize our field, but, in my opinion, the future of developmental biology lies in the inevitable integration of the great trends that saw their beginnings in the last two decades: it will feature extensive use of next-generation sequencing and bioinformatics to elucidate pre-translational 'omics while finding the etiology of developmental diseases and refining the protocols for regenerative medicine. This may well be considered a known and fixed future, but its predictability is not one that precedes boredom. For me, this is a future that is as exciting as any other development that anyone could imagine.

#### References

- Bryant, Z., Subrahmanyam, L., Tworoger, M., LaTray, L., Liu, C. R., Li, M. J., van den Engh, G. and Ruohola-Baker, H. (1999). Characterization of differentially expressed genes in purified *Drosophila* follicle cells: toward a general strategy for cell type-specific developmental analysis. *Proc. Natl. Acad. Sci. USA* **96**, 5559-5564.
- Chen, P. Y., Mannings, H., Slanchev, K., Chien, M., Russo, J. J., Ju, J., Sheridan, R., John, B., Marks, D. S., Gaidatzis, D. et al. (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev.* **19**, 1288-1293.
- Fox, J. L. (1999). Complaints raised over restricted microarray access. *Nat. Biotechnol.* **17**, 325-326.
- Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., Takayama, M., Asada, K., Mirochnitchenko, O., Inouye, M. et al. (2006). The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res.* **34**, 1765-1771.
- Roux, W. (1888). Beiträge zur Entwicklungsmechanik des Embryo. V. Ueber die künstliche Hervorbringung 'halber' Embryonen durch Zerstörung einer der beiden ersten Furchungszellen, sowie über die Nachentwicklung (Postgeneration) der fehlenden Körperhälfte. *Virchows Arch. Pathol. Anat. Physiol.* **114**, 113-153; 246-291.
- Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-470.
- White, K. P., Rifkin, S. A., Hurban, P. and Hogness, D. S. (1999). Microarray analysis of *Drosophila* development during metamorphosis. *Science* **286**, 2179-2184.