

The ‘omics revolution: how an obsession with compiling lists is threatening the ancient art of experimental design.

Claudio D Stern

Department of Cell & Developmental Biology, University College London, Gower Street, London WC1E 6BT, UK. c.stern@ucl.ac.uk

The last quarter of a century ushered in a transformation in Biology following the huge successes in sequencing entire genomes of organisms (including humans) with increasing precision and speed and rapidly decreasing cost: genomics. Later other “omics” approaches followed: proteomics, metabolomics, epigenomics, and many more. Huge quantities of data were generated and new computational methods to find hidden correlations emerged. Almost imperceptibly, however, it seems that this has led to a crisis in our ability to design elegant experiments to establish causality.

Imagine an extra-terrestrial excursion to planet Earth to discover about life on this planet. A spaceship with alien scientists lands in a park with many statues on display, representing important people from British history. The aliens unload their sophisticated analytical equipment and start to investigate. The first one uses a mass spectrometer – s/he approaches each statue and analyse it: one statue contains 5% copper, 10% iron, 17% aluminium, 63% lead. Another statue seems to consist mainly of carbon and a little water. A third object is made of granite. And so on. Another scientist follows: their analysis with a powerful super-resolution electron microscope reveals different sets of atomic structures. A third scientist uses a balance and lists the mass of each statue. A fourth measures the dimensions – some statues are very small, others quite large, and they have different overall shapes: some are tall, some are rather rounded. A fifth scientist examines the spectrum of light reflected from each, which suggests that they may have different colours. And so on. In all, the team of experts concludes that earth contains various objects of various dimensions, composition, etc. and then meet to speculate on their possible functions, purposes, and so on based on their observations. Perhaps the next day they might have embarked on performing “functional” experiments to test whether particular elements are “necessary” for the object by removing one element (carbon, lead, etc.) at a time. Of course, even with all their sophisticated approaches, they are unable to conclude that one of the statues depicts Winston Churchill standing tall, holding a cigar, another one is a bust of Queen Elizabeth II, a third represents King Arthur ...

Despite the far-fetched analogy, this is not so different from how many scientific studies currently approach developing an understanding of the biological world. In the last few decades we seem to have developed an obsession with generating lists: of gene expression, of physical parameters, etc. The lists are not only of components and other attributes, but might also include an exhaustive catalogue of which of these might be “necessary” and “sufficient” for whatever process is being studied. Will these approaches help us to understand the *rules* that govern the behaviour of the biological systems under study? Perhaps these comprehensive databases are as far from this level of understanding as the extra-terrestrials uncovering the true meaning and purpose of the statues in the park.

What is missing from both approaches? Higher resolution, greater precision, more detail, more sophisticated techniques are clearly not the answer. Almost the opposite, our understanding of both situations almost diminishes in proportion to the amount of detail and precision. Clearly what is missing in both cases is standing back at a distance, asking the right questions and understanding at what *level* should those questions be aimed. This series of essays started from the question of how

to establish *causality* in Biology, but even this concept does not quite convey the true depth of the problem, because Genetics (crudely, listing which genes and gene products are “necessary” for some property of the system being studied) is still unable, on its own, to generate this level of understanding. One point of view argued by our colleagues is that mathematical or computational *modelling* can help. This approach can provide a test for whether our representation of the system is sufficient to mimic its observed behaviour, but it still requires an appropriate level of enquiry, and an overall hypothesis, in addition to the ability to test the model (which after all is only an analogy) rigorously in the real world. The challenge, as for the extra-terrestrials, is how to identify the right questions.

I am particularly partial to the discipline of “Experimental embryology” (the usual modern translation of the approach first articulated by Wilhelm Roux in the 1890s as *Entwicklungsmechanik* – the mechanics/mechanisms of development). This discipline aims to uncover general rules about the behaviour of a system even if its component parts are not yet known. The idea arose mainly from the debate from experiments being done by Roux and by Hans Driesch around that time, where the former had killed one of the first two blastomeres after the first cell division with a hot needle (obtaining a half-gastrula) and the latter (and later Hans Spemann) had approached an embryo of a similar stage by separating the first two cells (obtaining two half-sized, but complete embryos). The difference between the two results and the manipulations that had generated them points to cell contact, in the context of which, even the remnants of a membrane proved to be important; however, it does not reveal the identity of the membrane components, or how the contact conveys information to the remaining cell about whether it should develop as a whole or as a half-embryo. Unfortunately, it is often difficult to design experiments like this, but the principle remains.

Apart from Experimental Embryology, classical Genetics has some of the same attributes. Before the discovery that DNA is the main genetic material, the principles of inheritance had been uncovered in great detail and “genes” described before understanding their physical nature or how they act; as early as 1911, Alfred Sturtevant and his teacher Thomas Hunt Morgan had even managed to predict the “distance” between genes according to the frequency of recombination between them. Similarly, Physiology made huge advances, for example by elegant mathematical descriptions of the action potential, long before its physical basis was understood. All of these approaches have in common that they can, in principle, make accurate predictions of how the system will behave even before there is any understanding of the true physical “mechanism” or even of all of its components. Just as in Physics it is possible to propose a simple equation for gravitational force that provides surprisingly accurate predictions of the behaviour of two objects based merely on their respective masses and the distance between them, as Newton did. Physicists do not seem to care so much about uncovering the physical nature of the forces themselves – the predictive equation is a sufficient explanation. So why have biologists, in general, become so obsessed with listing components in ever increasing detail?

Some current trendy approaches and problems

Single-cell mRNA sequencing (scRNAseq) for cell lineage tree reconstruction

Huge technical advances in molecular biology have recently opened the exciting opportunity of sequencing all mRNAs present in a single cell. When used as part of an appropriately designed experiment, this is an extremely powerful methodology. But there are many examples in the recent literature that use this methodology to draw conclusions which, in my opinion, exceed what this

information can reveal – for example, conclusions about cell lineage relationships during development, reconstructed solely from scRNAseq data.

There are several reasons why this can be a problem. During development, cells move around in the embryo. As they do so, they enter and leave territories characterised by gene expression domains. For example, during gastrulation, cells enter and leave the tip of the primitive streak (the node), which is “marked” by expression of node-specific genes like the transcription factor Goosecoid. Before they enter the node, cells do not express this gene (regardless of whether they are destined to express it later or not). After they leave the node, cells no longer express it, regardless of their destination and regardless of whether or not they have passed through the node to reach that destination. Although a few genes can be exclusive to a particular terminal fate, a more appropriate rule for developmental gene expression is that *genes do not mark cell fates but rather mark cell states*: they are expressed according to what the cell is doing at that moment. Furthermore, some genes undergo extremely fast dynamic changes in their expression, for example during the process of somite formation, where a number of genes turn on and off cyclically with a period of about 90 min. And there are examples of genes that change expression even faster. Since RNAseq can only be done once for any one cell, every cell is unique and cannot be followed in time. Because of cell heterogeneity (so beautifully demonstrated by scRNAseq, but also by more classical techniques such as in situ hybridisation, which often reveal spatial mosaicism in gene expression), each sample is rather unique. More importantly, perhaps, due to cell movements and other reorganisations in the embryo, sampling a particular embryonic location at different times in development may not follow the same cells, because some cells will be entering and leaving that location, as in the above example of the node at successive stages: some cells have left, others entered, and only a very small number will be present in both samples. Therefore, reconstructing a “lineage” from scRNAseq information alone would be analogous to our extra-terrestrial colleagues trying to unravel the rules of football by taking a few photographs during a number of matches in the main stadium. No matter how many matches they sample, it is impossible to reconstruct any of the games or, of course, the rules. And as with scRNAseq in development, different players are engaged in each match, even in the same stadium.

scRNAseq can indeed be used to illuminate processes that cells undergo during a cell fate transition, but at best, when used alone, it is only predictive/suggestive. To unravel a hierarchical set of decisions or a lineage tree requires cell labelling and following the cells over time. At present it is still technically impossible to follow changes in the expression of multiple genes in a single cell with the required time resolution. scRNAseq generates lists of combinations of genes that are expressed together in individual cells, but without greater knowledge of the specific state of those cells and the conditions of the experiment -- as well as the relationship to a real time line -- the interpretation of these lists must be undertaken with caution. In particular, it must be done with more underlying biological knowledge if it is to serve as more than a suggestion.

Mechanobiology

In the 1970-1980s, many cell and developmental biologists became interested in the possibility of non-genetic events that could influence cell behaviours. Cell adhesion, the physical and mechanical properties of tissues, electrical (ion) currents flowing between cells and even outside a tissue to generate extracellular potentials, and even electromagnetic phenomena were measured and sometimes functionally explored to determine whether they could act as “epigenetic” mechanisms, alongside gene expression, to regulate cell behaviours and developmental decisions. It is somewhat

odd that, not only did this field largely die in the 1990s (probably as “transcriptomics” became a more widespread approach, alongside the sequencing of many genomes), but even more so that it has recently made a major comeback. It is extraordinary to see how many international meetings feature sessions on this topic, but even more so to see how much of the early work is ignored, and old experiments often repeated. Generally, the studies involve measuring something: forces, fluidity, strain, stress, voltage, etc. in a tissue, perhaps at different times. This work also generates lists. The modern studies take advantage of improved methods with much greater precision, and therefore measure the parameters with more decimal places. With only a few notable (and beautiful) exceptions, however, these studies also largely fall short of establishing true causality for those forces. Experimental perturbation can help to address some of this, but it is very difficult to avoid introducing other variables that are not part of the hypothesis being tested. Therefore, these studies often rely on correlations between measurements and behaviours, rather than establishing clear causal relationships. There are exceptions of course, and some new findings that have been particularly exciting often include demonstrations of the connections between physical parameters and changes in gene activity – thus, we have found out that cells can sense the radius of curvature of one of their surfaces and change gene expression accordingly, and also that there are specialised proteins (such as the PIEZO proteins) that can respond differently to mechanical forces by modulating the activity of an ion channel and thus generate an electrical signal; moreover, these are expressed even in tissues that are not generally considered to be mechanosensitive or excitable. In this case, however, the main stumbling block is probably that methods to manipulate these physical parameters within the correct, biologically-relevant range, and without affecting other process as a side-effect, have not yet been developed sufficiently. But another important underlying problem is that we have not yet started to understand “where” the code that governs how an embryo or its component parts develop resides: where is the “computer programme”? How can evolution act on this type of determinant, and how is it inherited?

Genetics

Even classical genetics tends to generate lists – in this case, of phenotypes associated with particular genetic changes. This has been hugely transformative in the whole of Biology – most of the genes/proteins now known to play important roles in development of all metazoans have been discovered through forward genetic approaches in *Drosophila*, bacteria and yeast (later coupled by mapping and identification of the DNA sequences of the genes and the specific molecular consequences of the mutation). Remarkably, most proteins that play roles in *Drosophila* are hugely conserved, as are the “pathways” that connect subsets of these proteins to define regulatory relationships. These pathways have turned out to work as “toolkits”, used for many different processes for which they have been co-opted, and which may not themselves be very conserved. They are therefore also often used multiple times in development. Thus, evolution has acted strongly to conserve the pathways, but has allowed these to be “called” (like a subroutine/procedure in a computer program) for very different purposes. Therefore, living forms have diverged from each other by deploying the pathways at different times and places.

But moving beyond this, and using genetics alone to elucidate how a particular process takes place, has been much more difficult. The reason is probably that the experiments performed tend to be based on “loss-of-function” via abrogation of protein function or gene expression (even if “conditional” to a particular time and place). This can only reveal the processes for which the protein is essential. Alone, it is much more difficult to establish the mechanism. So efforts to generate large

numbers of knockout mice and a catalogue of phenotypes also generate lists. The resulting mutants are good resources for other experiments, of course.

I was privileged in 1974, as an undergraduate, to attend a meeting at the Royal Society where Sydney Brenner gave a talk introducing the nematode as a model system for studying development. He used no slides, just a blackboard and chalk. With a few simple diagrams and numbers, he announced that because of the relatively small number of cells and small genome, we should be able to determine all cell lineage relationships as well as the functions of all genes by using genetics. He predicted that once we knew all of these relationships and all of these functions, we would really understand development. On the bus trip back after the meeting, and long into the night afterwards, my classmates and I discussed whether Brenner might turn out to be right. We now know that he was certainly correct that it would be feasible within a comparatively short time to elucidate both the complete lineage of all cells of this organism as well as to compile the consequences of loss of function of all its genes. But he was certainly wrong in his prediction that when this was achieved, we would understand development, either in this organism or in any other. Why? Because again, all that these approaches generate is *lists*: in this case, of genealogical relationships between cells, and of the consequences of loss of function of every gene. But the rules do not emerge naturally or spontaneously from such lists, even if we continue to compile more lists.

'Omics

Of course, the most extreme example of list-making is the multitude of 'omics approaches. This approach emerged in the 1990s, initially from efforts to sequence the whole human genome (and those of other organisms), and as the technology progressed rapidly, it became feasible to do this increasingly efficiently and with increasing precision. Genomics was followed by Transcriptomics, Proteomics, Lipidomics, Glycomics, Metabolomics, Epigenomics, Metagenomics, Functional genomics (systematically generated, comprehensive mutant collections), ... The volume of data generated by these approaches multiplied rapidly across the whole of biology. Alongside, improvements in data storage, computational power and many software resources have facilitated the analysis of these massive datasets. At one point, a British funding body, the BBSRC, decided as a strategic priority to focus on "data-driven science": approaching the mass of data with computational approaches to find commonalities, and even to identify the questions as well as the answers. As with the examples above, these will only be mere correlations, even if they are strong ones. But correlation does not imply causation. Establishing causality requires some form of experimental design, which must be hypothesis-driven; it cannot rely entirely on compiling yet more lists and presenting them to a computer to discover hidden patterns.

The various 'omics are powerful techniques, and the resources they generate have been transformative. But few could have predicted that in just two decades this methodology would become so pervasive as to create a whole philosophical approach, almost replacing experimental design. Listing everything seems to have become the default method of approaching any biological problem.

Some of my colleagues (authors of some of the accompanying articles) advocate that either a more quantitative approach, or modelling, provide the answer. But even "Quantitative Biology" is an example of 'omics: it seems to advocate measuring everything/anything with ever increasing precision and compiling lists of the parameters. Likewise, "modelling" can be useful to make predictions, and it can certainly be fun, but at best it can only mimic the more complex system by

asking whether it is possible, in principle, to explain the phenomenology by a simple subset of chosen interactions. 'omics provides correlations; modelling provides analogies – neither can establish causality.

The way forward?

“Good” science is more than discovering new knowledge at the fastest possible rate: the elegance of the experimental design, and of the demonstrations that establish the knowledge, are at least as important: it isn't just **what** you discover, but **how** you show it that matters – scientific-aesthetics. Not so long ago the top-rated journals understood this and used it as an important criterion to determine whether the paper was “Nature-/Cell-/Science-worthy” or whether it belonged in a “more specialised journal”. The high-ranking journals also seem to have lost their way, and apart from tabloid-like headline-making papers, they also publish rather a lot of lists.

I am a little worried that the current generations of biologists are increasingly exposed to list-making as the only approach to discovery. Undergraduates seem to find it very difficult to design experiments for a purpose, perhaps in part because their teachers themselves have lost the art? Many grant applications coming through panels seem to rely on the same techniques or strategies: make a list, pick one element and do “functional experiments” by knocking out and/or overexpressing (or the equivalent for other types of lists). This will do nothing more than contribute to yet more lists – not deeper understanding of a process.

However, I am optimistic that we will eventually move away from this obsession, that we will re-discover the beautiful classical art of elegant experimental design. As has always been the case, this requires identifying a question (for example a biological process that we do not understand), thinking through the possible (and plausible) mechanisms that *could* explain this process, and testing each one systematically. Good experimental design should aim to reduce the number of variables systematically; at its best, this should lead to a compelling demonstration of causality. But even when such final proof impossible, there is considerable intrinsic beauty in hypothesis-driven, careful experimental design.

Acknowledgements: I am grateful to Mike Levin both for inviting me to participate in the original multi-author contribution (Bizzarri et al., Nature Reviews Cell Biology 20: 261-262, 2019) and for his helpful comments on the present manuscript.