nature REVIEWS GENETICS

From OMICS to systems biology

Marc Vidal and Eileen E. M. Furlong

The field of 'omics' currently polarizes the community of biologists. What do all these terms mean? Is there a fundamental difference between functional genomics and genetics? Is proteomics not what we used to call biochemistry?

The transcriptome, proteome, interactome and metabolome of a cell correspond to tangible entities (see glossary). Transcriptomics, proteomics, interactomics and metabolomics are fields of studies that focus on the properties of those entities. They should not be just buzzwords, mostly used to attract large sums of funding, public or private. Instead, such neologisms hopefully help to formalize and address biologically relevant questions.

Suffixes and neologisms are not new in science. They reflect the evolution of our collective thoughts. 'Omes' were already used in the beginning of the last century to indicate the 'wholeness' of biological systems, such as in 'biome' or even 'rhizome'[1]. Other suffixes such as 'ion' and 'on' help us to conceptualize the most elementary particles, such as in 'ion, cation and anion', or in 'photon, electron, proton and meson'.

Considering the wholeness of cellular systems by conceptualizing the proteome or the metabolome might be as important for our understanding of biology as, for example, the wholeness of economical markets is to the study of macroeconomics.

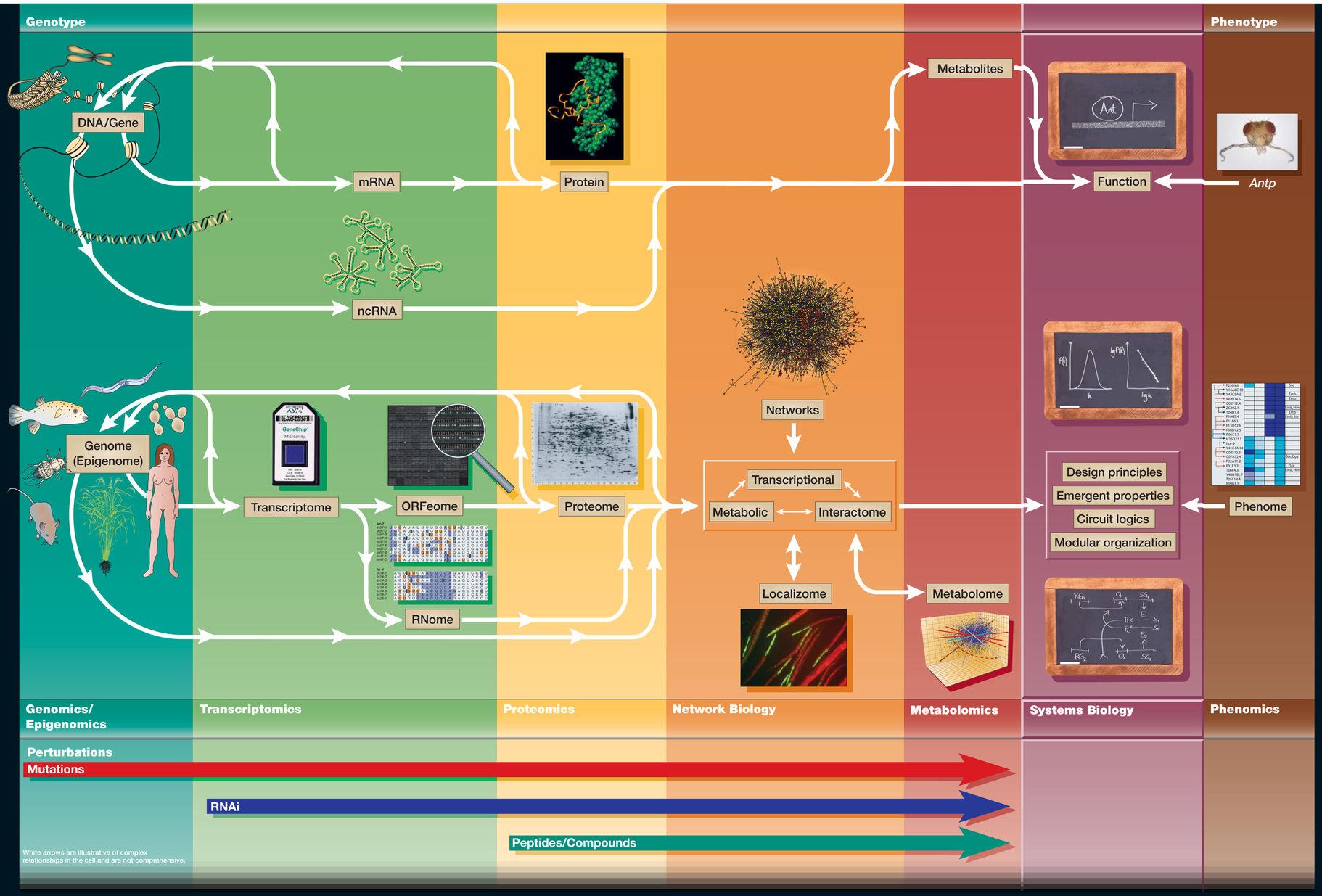
Glossary

Genome, a term first used in the 1920s[1], is now accepted as the whole complement of genes of an organism. Likewise, the concept of proteome, coined in the mid-1990s, helps us to conceptualize the full complement of proteins contained in an organelle, cell, tissue, organ or organism. Proteomics studies global and local properties of the proteome.

Other omes also correspond to tangible entities: ORFeome is the set of protein-encoding open reading frames (ORFs); transcriptome refers to all coding transcripts; RNome to all non-coding RNAs (ncRNAs) and metabolome is the full complement of metabolites (metabolomics and metabonomics refer to nearly complete sets of metabolites in cells and organisms, respectively).

Perhaps less tangible, but equally important[2], are: the interactome, the complete set of macromolecular interactions (protein–protein, DNA–protein, RNA–protein) in a cell, tissue, organ, or a whole organism; the localizome, a dynamic concept that summarizes where and when all gene products are located in a cell throughout its life or genes are expressed during the organism's development; the phenome, information collected in a standardized and systematic manner on large numbers of phenotypes observed upon perturbation of nearly all genes or gene products of an organism.

By interacting with each other, macromolecules form complex networks organized into systems with properties that extend beyond the function of each individual molecule (emergent properties). Systems biology attempts to provide predictive models of the behaviour of such molecular systems (design principles). Integrated omics analyses can help to model cellular networks and to identify, for a given biological process, or 'module'[3], the network components and their functional interactions (modular organization). Inside modules, small numbers of molecules can form circuits that have their own functions. An important goal of systems biology is to study such 'circuit logics'.



References

1. Lederberg J. & McCray A.T. 'Ome Sweet 'Omics – a genealogical treasury of words. *The Scientist* 15 August (2001).
2. Vidal M. A biological atlas of functional maps. *Cell* **104**, 333–339 (2001).
3. Hartwell L. *et al* From molecular to modular biology. *Nature* **402** (6761 Suppl.), C47–52 (1999).

Antp image courtesy of Walter Gehring. Localizome image courtesy of Denis Dupuy. ANTP protein image courtesy of Biomolecules Gallery, RIKEN Tsukuba Institute, Japan.

Since Mendel's wrinkled peas and Morgan's white-eyed fruitflies, the main framework of genetics has rested upon basic questions such as: 'How do different genotypes give rise to different phenotypes'? Genetics identifies the molecular players of cellular biology, usually one by one, and especially when combined with biochemistry, can help to establish some of their *in vivo* functions. As one of the memorable examples, the discovery that *antennapedia* fly mutants develop legs in place of antennas ('phenotype') and that the corresponding wild-type gene product binds to and activates the promoter of 'downstream' target genes ('protein') led to a model whereby transcription factors can act as important developmental switches ('function'). More recently, the finding that loss-of-function mutations in *lin-4* and *let-7*, two genes that express non-coding RNAs (see 'ncRNA'), confer developmental phenotypes emphasized the biological importance of ncRNAs.

These approaches, based on the 'one-gene/ one-function' concept, have led to the discovery of an impressive list of molecular players and functions. Nonetheless, key questions remain unanswered. Take cellular differentiation: the function of individual transcription factors can provide molecular descriptions of cellular phenotypes as they change during development, but they usually fail to provide mechanistic explanations. How are those transcription factors themselves regulated? What are their target genes? And importantly, how can particular developmental switches be 'remembered' at the molecular level through many generations of asymmetrically dividing cells. In other words, how do we explain that different cells in a single organism have different 'cellular phenotypes', despite usually having the same genotype? To what extent can classical genetics answer 'how a single genotype can lead to different cellular phenotypes'?

The answers will probably emerge from studying the design principles and emergent properties of complex molecular systems formed by proteins and ncRNAs. It is to be hoped that both local and global modelling of the cellular networks will reveal the modular organization of cells and the logics of their molecular circuits.

Genome sequencing and annotation projects provide a first-degree approximation of the 'parts-lists' that compose such macromolecular networks. Systems biology attempts to piece the parts together. To understand the global relationships between genotype, environment and phenotype, we need to consider the metabolic, transcriptional and protein interaction networks (interactome) formed by nearly complete sets of genes (genome), transcripts (transcriptome), ncRNAs (RNome), proteins (proteome) and metabolites (metabolome).

Systematic perturbations of genes (mutations), transcripts (RNAi) or proteins (using peptides or compounds) should help with understanding complex cellular networks. The scoring of large numbers of standardized phenotypic descriptors followed by clustering analysis reveals features of the 'phenome' of an organism.

Poster design: Tom Wilson and Simon Fenwick.
Image concept: Philippe Lamesch.
Edited: Magdalena Skipper.

http://www.nature.com/reviews/genetics