Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

New genes contribute to genetic and phenotypic novelties in human evolution

Yong E Zhang[1] and Manyuan Long[2]

New genes in human genomes have been found relevant in evolution and biology of humans. It was conservatively estimated that the human genome encodes more than 300 human-specific genes and 1000 primate-specific genes. These new arrivals appear to be implicated in brain function and male reproduction. Surprisingly, increasing evidence indicates that they may also bring negative pleiotropic effects, while assuming various possible biological functions as sources of phenotypic novelties, suggesting a non-progressive route for functional evolution. Similar to these fixed new genes, polymorphic new genes were found to contribute to functional evolution within species, for example, with respect to digestion or disease resistance, revealing that new genes can acquire new or diverged functions in its initial stage as prototypic genes. These progresses have provided new opportunities to explore the genetic basis of human biology and human evolutionary history in a new dimension.

Addresses
[1] Key Laboratory of Zoological Systematics and Evolution & State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing, China
[2] Department of Ecology and Evolution, The University of Chicago, Chicago, USA

Corresponding authors: Zhang, Yong E (zhangyong@ioz.ac.cn) and Long, Manyuan (mlong@uchicago.edu)

Current Opinion in Genetics & Development 2014, **29**:90–96

This review comes from a themed issue on **Genetics of human origin**

Edited by **Aida Andes** and **Katja Nowick**

http://dx.doi.org/10.1016/j.gde.2014.08.013

Introduction

Evolutionarily new genes, referred to genes emerged in recent evolution [1], have attracted a broad interest, since the first mechanistic model was proposed in the 1930s [2]. Thanks to extensive studies of molecular evolution and genomic biology in the last decade, a dozen of distinct molecular mechanisms to generate new genes were found, including the most frequently investigated DNA-based or RNA-based duplication mechanisms and a recent additional hot topic of *de novo* origination [1,3]. These mechanisms lead to pervasive new gene origination, which in turn participated in lineage-specific or species-specific phenotypic evolution [4]. For human biology, tremendous efforts have been dedicated to study human-specific genes absent in other primates or polymorphic genes within human species, which leads to a significant progress in understanding how often these new genes contributed to phenotypic evolution and how they are implicated in disease [5,6•,7].

To discuss the progress in technical and conceptual investigation of new human genes, we provide here a concise and updated overview. We first focus on the rate and describe a few efforts in identifying primate-specific or even human-specific new genes encoded by the human genome. We describe the emerging themes in the functionality of these recently evolved genes and highlight their significance for brain and testis evolution. Then, we discuss a new hypothesis regarding the phenotypic evolution by new genes in the light of recent functional data indicating that new genes can promote tumorigenesis, while evolving advantageous functions. We further discuss the initial stage of new gene evolution when a new gene is polymorphic in a species population and discuss how these genes contribute to phenotypic difference between individuals or populations. We end the review with a summary of potentially important directions.

The human genome gains a high flux of new genes

The pioneering effort via cDNA array-based comparative genomic hybridization (aCGH) identified 134 genes showing copy number expansion after the split of human and great apes [8]. Further genomic analysis including three additional mammalian species identified 689 human-specific genes, that is, the ones not shared by chimpanzees, and 870 hominoid genes shared by human and chimpanzee but absent in mouse and dog [9]. A third analysis of 18 vertebrate genomes detected 389 human-specific genes and 1828 primate-specific genes [10]. Besides different identification strategies, the changing number could also result from ever-changing annotation. For new genes, this issue became more serious due to their poor conservation and narrow expression [11••]. For example, out of 1828 primate-specific genes, more than half were revised by later Ensembl updates as pseudogenes or noncoding transcripts, or unduly removed from the annotation [11••]. In other words, the annotation database is getting more conservative when including entries of new gene. Such an issue should be cautioned when studying new gene evolution.

The difficulty that the unstable and insufficient annotation brought to the study of new gene evolution was demonstrated by the contrasting number of human-specific *de novo* genes across different studies. Comparative analyses across multiple primate genomes in Ensembl v47 revealed three human-specific *de novo* genes supported by both transcription and proteomics data [12]. Pooling of multiple Ensembl versions (v40–v56) led to an exciting discovery of 60 human-specific *de novo* genes [13]. A third analysis based on Ensembl v51 pooled out 11 human-specific *de novo* genes [14]. All these efforts are similar technically: (1) to call proteins with the corresponding orthologous region in outgroups incapable of coding the open reading frame in the genomes of recent human ancestors; (2) to ensure that candidate *de novo* genes are supported by peptide databases. However, as pointed out in [15], the difficulty roots in the lability of human annotation of new genes and the arbitrariness of bioinformatic parameters. Nevertheless, combining complementary efforts on both duplicated new genes and *de novo* new genes, it seems prudent to conclude that substantial changes occurred in the human gene reservoir with about 300 human-specific genes and 1000 primate-specific genes added.

Primate-specific or human-specific new genes are often implicated in brain development and male reproduction

Whether or not a new gene contributes a crucial phenotypic effect in evolution is an interesting problem. As one of the early reported primate-specific gene families, *morpheus* was found to encode nuclear pore complex interacting protein (NPIP) with wide transcription in numerous tissues and organs [16] and its specific function has been known more for its activity involved in the HIV replication [17]. Recently, numerous cases of new genes were reported related to various molecular functions or phenotypic effects (more examples can be seen in Table 1). Quite a few cases appear to be related with brain functions such as the glutamate dehydrogenase 2 (*GLUD2*) [18,19] or the neuroblastoma breakpoint (DUF1220) family [20,21]. A recently well characterized case in support of the significance of new gene emergence for human brain evolution is Slit-Robo Rho GTPase-activating protein 2c or *SRGAP2C*, which is a DNA-level duplicate originated around 2 million years ago [22••]. As a partial copy, *SRGAP2C* inhibits the function of its parental gene *SRGAP2A* and induces neoteny during dendric spine maturation [23].

The enriched recruitment of new genes into brain expression is not only detected by these case studies, but also strongly supported by genome-wide studies. Comparative transcriptome profiling across major organs revealed that the proportion of brain transcriptome contributed by primate-specific genes in human is significantly higher than that contributed by rodent-specific genes in mouse [31].

Analogously, transcriptome profiling of hominoid-specific and human-specific *de novo* genes also showed that these genes tend to be transcribed in brain and testis [13,14]. Primate-specific genes transcribed in brain is enriched for zinc finger (ZNF) genes [31], which appear to be mainly contributed by the *Kruppel*-type or KRAB family [32]. Interestingly, about 40% of primate-specific KRAB-ZNF genes are differentially transcribed between human and chimpanzee prefrontal cortex, which may lead to extensive gene expression difference between the two species [33]. Why brain acts like an evolutionary hotbed in recruiting new genes likely roots in the complexity of its molecular network. Before the genomic era, it was

Table 1

Examples of new genes evolved after the split of primate from other mammals. Human-specific new genes refer to those genes absent in the other primates including chimpanzee. Homininae-specific genes refer to those shared by human, chimpanzee and gorilla. Hominoid-specific genes refer to those evolved recently in the lineages of apes but absent in rhesus monkey and other primates. Primate-specific genes refer to those absent in non-primate mammals.

Gene	Origination mechanism	Age	Function	Citation
Brain-related new genes				
GLUD2	RNA-based duplication (retroposition)	Hominoid-specific	Glutamate metabolism in brain	[18,19]
DUF1220 family	DNA-based duplication	Primate-specific	Transcribed in brain	[20,21]
SRGAP2C	DNA-based duplication	Human-specific	Dendric spine maturation	[22••,23]
Testis related				
SPANXA/D	DNA-based duplication	Homininae-specific	Spermatid morphogenesis	[24]
Cancer related				
CT45A1	DNA-based duplication	Primate-specific	Upregulate oncogenic and metastatic genes	[25]
TBC1D3	DNA-based duplication	Hominoid-specific	Modulator of epidermal growth factor receptor signaling pathway	[26,27]
NCYM	*De novo*	Homininae-specific	Stabilize the oncogene *MYCN*	[28••,29]
PBOV1	*De novo*	Human-specific	Possibly repress tumorigenesis	[30]
Other				
Morpheus family	DNA-based duplication	Primate-specific	Broadly transcribed	[16,17]