Genesis and Functional Relevance of Orphan Genes in Evolutionary Biology

Sainath K[1]*, Praveen K[3], Akshaya M[2], and B. P. Maruthi Prasad[4]

[1]*Department of Genetics and Plant Breeding, UAS, Raichur*
[2]*Department of Genetics and Plant Breeding, UAS, Dharwad*
[3]*Department of Agricultural Microbiology, UAS, Bangalore*
[4]*Department of Genetics and Plant Breeding, UAS, Bangalore*

Corresponding Author
Sainath K
Email: sainathkarbhari98@gmail.com

OPEN ACCESS

Keywords

Orphan gene, Phylostratigraphy, BLAST, Vertical gene transfer, Retro position

How to cite this article:

Sainath, K., Praveen, K., Akshaya, M. and Prasad, B. P. M. 2025. Genesis and Functional Relevance of Orphan Genes in Evolutionary Biology. *Vigyan Varta* 6 (12): 7-10.

ABSTRACT

Orphan genes, also known as taxonomically restricted genes (TRGs), represent species-specific coding sequences with no detectable homologues in other organisms. These genes are key drivers of evolutionary innovation and adaptation, often arising through mechanisms such as gene duplication, divergence, fusion, fission, horizontal gene transfer and retro position. While many orphan genes originate de novo from non-coding regions, others evolve beyond recognizable similarity to ancestral genes. The concept of orphan genes was first introduced during the yeast genome sequencing project in 1996, where they accounted for about 26% of the genome. Subsequent research established their widespread presence across taxa and highlighted their roles in lineage-specific traits and functional diversification. Identification of orphan genes primarily relies on computational approaches such as BLAST and phylostratigraphy, which detect sequence homology and estimate gene age, respectively. Despite challenges in detecting short or rapidly evolving sequences, orphan genes continue to provide valuable insights into genome evolution, species diversification, and the emergence of novel biological functions.

INTRODUCTION

Orphan genes are defined as genes with coding sequences utterly unique to the species. Genes without detectable sequence similarity in the genomes of other organisms (Tautz *et al.,* 2011). Genes are generally classified as being orphans if they lack coding sequence similarity outside their species. They are also called as "taxonomically restricted genes" (lineage specific). Genes born from non-genic sequence, as well as descendants of ancient genes whose coding sequences have changed beyond recognition.

NOVAL GENE FORMATION FROM ANCESTRAL GENES MECHANISMS

Duplication: This occurs when a segment of DNA is copied, resulting in two identical genes (or gene copies). Gene duplication can happen through various mechanisms, such as unequal crossing over during meiosis or replication errors (Otto and Yong, 2002).

Divergence: After duplication, the two gene copies can accumulate mutations independently. Over time, these mutations can lead to differences in the gene's function, expression or regulation. This process allows one gene copy to potentially take on a new function (neofunctionalization) or for both copies to retain similar functions but diverge in their expression patterns (Taylor and Raes, 2004).

Gene fusion: Gene fusion is the process by which two or more separate genes combine to form a single gene, resulting in the production of a fusion protein. This can occur through various mechanisms, such as chromosomal rearrangements or translocations, and can lead to new functions or regulatory properties that may provide evolutionary advantages (Annala *et al.,* 2013).

Gene fission: Gene fission is the process by which a single gene splits into two or more separate genes. This can occur due to mechanisms such as chromosomal breakage, duplication followed by divergence, or the insertion of transposable elements. Gene fission can lead to the specialization of gene functions and contributes to genetic diversity and evolutionary adaptation (Snel *et al.,* 2000).

Horizontal gene transfer (HGT): Horizontal gene transfer is the process by which genetic material is transferred between organisms in a manner other than traditional reproduction (vertical gene transfer). This occurs primarily in prokaryotes, such as bacteria, but can also happen in eukaryotes. HGT can occur through several mechanisms, including:

1. **Transformation**: Uptake of free DNA from the environment.
2. **Transduction**: Transfer of DNA between bacteria via bacteriophages (viruses that infect bacteria).
3. **Conjugation**: Direct transfer of DNA between two bacteria through cell-to-cell contact.

Retro position: Retro position is the process by which an RNA molecule is reverse-transcribed into complementary DNA (cDNA) and then integrated into a genome. This mechanism can lead to the creation of new genes or gene copies, contributing to genetic diversity and evolution. (Hughes, S.H., 2015)

HISTORY OF ORPHAN GENES

The term "orphan gene" was first introduced by Bernard Dujon during the yeast genome sequencing project in 1996, where such genes accounted for about 26% of the yeast genome. Initially thought to lack homologues due to limited sequencing data, orphan genes were later found to persist across all species, establishing their ubiquity in genomes.

Typically comprising 10–30% of genes, they became the focus of evolutionary studies after 2000. Research in *Caenorhabditis briggsae* (2003) suggested that orphans evolve too rapidly to be detected, while studies in bacteria (2005) confirmed their authenticity and role in adaptation. The term "taxonomically restricted genes (TRGs)" was later proposed to describe them more precisely. Functional evidence followed with the discovery of BSC4 in yeast (2008), a gene evolved de novo from non-coding DNA, and QQS in *Arabidopsis thaliana* (2009), which regulates plant composition through interaction with a conserved transcription factor. Genes are classified as orphans when no orthologues are found in related species, commonly identified using the Basic Local Alignment Search Tool (BLAST). However, short and rapidly evolving genes may escape detection, highlighting both the evolutionary novelty and functional significance of orphan genes (Seçkin *et al.,* 2025).

IDENTIFYING ORPHAN GENES

1. BLAST
2. Phylostratigraphy

BLAST (Basic Local Alignment Search Tool)

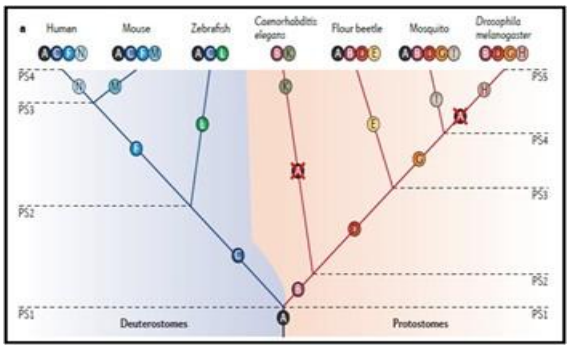
BLAST is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of proteins or the nucleotides of DNA and RNA sequences. BLAST finds regions of local similarity between sequence. It is the preferred method for detecting homologues of gene in other species. Most sensitive tool for detecting even such remote homologues Position – Specific Iterated BLAST (PSI- BLAST). (Samal *et al.,* 2021).

Types of BLAST:

- **BLASTN**: Compares nucleotide sequences to nucleotide databases.
- **BLASTP**: Compares protein sequences to protein databases.
- **BLASTX**: Translates nucleotide sequences into proteins and compares them to protein databases.
- **TBLASTN**: Compares protein sequences to translated nucleotide databases.
- **TBLASTX**: Compares the translated forms of nucleotide sequences against each other.

Phylostratigraphy:

Phylostratigraphy can be used to systematically identify all orphan genes with in the evolutionary lineages. Genes can be stratified by age *via* a technique known as Phylostratigraphy that traces modern genes back to their orphan founders. (Neme Garrido, R.T., 2011)



APPLICATIONS

- Evolutionary Studies: Phylostratigraphy provides insights into the evolution of

specific traits and functions by tracing the origins of related genes.

- Functional Genomics: It helps in understanding the evolution of gene function and the roles that different genes play in various organisms.
- Comparative Genomics: By comparing gene presence across different species, researchers can make inferences about common ancestors and evolutionary pressures.

CONCLUSION

Orphan genes represent a valuable source of genetic novelty and play a crucial role in species-specific adaptation and evolution. Their unique functions, often linked to stress tolerance, development and reproductive traits, make them important targets for crop genetic improvement. Understanding the origin, expression and function of orphan genes can help breeders identify new genes associated with desirable traits and develop improved varieties through molecular breeding and biotechnological approaches. Therefore, integrating orphan gene research into breeding programs offers great potential for enhancing productivity, resilience and adaptability in crop specie

REFERENCES

Annala, M.J., Parker, B.C., Zhang, W. and Nykter, M., 2013. Fusion genes and their discovery using high throughput sequencing. Cancer letters, 340(2): 192-200.

Hughes, S.H., 2015. Reverse transcription of retroviruses and LTR retrotransposons. *Mobile DNA III*, 20: 1051-1077.

Neme Garrido, R.T., 2011. *Phylostratigraphic analyses of mouse tissue transcriptomes and comparative genomics of orphan genes* (Doctoral dissertation, Georg August University Göttingen).

Otto, S.P. and Yong, P., 2002. The evolution of gene duplicates. *Advances in genetics*, *46*: 451-483.

Samal, K.C., Sahoo, J.P., Behera, L. and Dash, T., 2021. Understanding the BLAST (Basic Local Alignment Search Tool) program and a step-by-step guide for its use in life science research. *Bhartiya Krishi Anusandhan Patrika*, *36*(1): 55-61.

Seçkin, E., Colinet, D., Sarti, E. and Danchin, E.G., 2025. Orphan and de novo Genes in Fungi and Animals: Identification, Origins and Functions.

Snel, B., Bork, P. and Huynen, M., 2000. Genome evolution: gene fusion versus gene fission. *Trends in genetics*, *16*(1): 9-11.

Tautz, D. and Domazet-Lošo, T., 2011. The evolutionary origin of orphan genes. *Nature Reviews Genetics*, *12*(10): 692-702.

Taylor, J.S. and Raes, J., 2004. Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.*, *38*(1): 615-643.