Nomenclature Tutorial

A Tutorial for Proposing Zebrafish Gene Nomenclature

By the Zebrafish Nomenclature Committee
Feel free to contact us at: nomenclature@zfin.org

Finding an appropriate gene name and symbol to use for your newly discovered gene can be a frustrating endeavor! There are so many factors to consider, like whether homologs or orthologs are known in other species, whether it is a member of a gene family, and whether it has a duplicate. We hope the following tips will guide you and make the experience of gene naming more pleasant.

How to propose a good gene name and symbol:

1. Become familiar with the zebrafish nomenclature guidelines.

2. Search ZFIN (http://zfin.org/) to ensure that your gene hasn't already been named. A good way to do this is to use ZFIN BLAST. Use the Nucleotide - Nucleotide BLAST.
If there is no identical match check to see if your gene has a closely related gene such as genes created by the genome duplication or a tandemly duplicated gene.

Gene duplicates, resulting from the whole genome duplication, often have differences in the third nucleotide or "wobble" position in a codon. Closely related sequences with mismatches separated by stretches of 2, 5, or 8 identical nucleotides may indicate that the genes are duplicates.

3. If you believe your gene is the duplicate of a named zebrafish gene based on a high sequence identity and conserved synteny, propose the gene symbol comprised of the existing, named gene symbol with an "a" or "b" attached to the end. The existing gene will be renamed by adding an "a" or "b" to the original symbol depending on chromosomal location.

If you are reporting one or both duplicates and have mapped them to duplicate chromosome segments, make an effort to keep the "a" and "b" terminology consistent with other duplicates already mapped on the chromosomal segment.

> **Example**: You identify the two zebrafish duplicates of the human gene ZZZ. You have mapped one duplicate to LG3 near the hoxb "a" genes and the other duplicate you have mapped to LG12 near the hoxb "b" genes. Propose zzza for the gene on LG3 and zzzb for the gene on LG12 to conform to the nomenclature of hoxb genes already mapped to those chromosomes.

4. If the gene is new to zebrafish but has a human or mouse ortholog, check the Human Nomenclature Database or the Mouse Genome Informatics resources for the name and symbol used in those species:

Human: http://www.genenames.org/

Mouse: http://www.informatics.jax.org/

Human and mouse gene names in these two resources are approved by nomenclature committees.

5. If there is an established gene name and symbol in human for the ortholog, use the same symbol for zebrafish (of course following zebrafish nomenclature guidelines like all lower case italics, no "z" prefix, etc.). If there is no orthologous human gene, but there is an orthologous mouse gene, use the same name and symbol used for the mouse ortholog.
Example: Suppose you isolate the zebrafish ortholog of the human "dihydrolipoamide branched chain transacylase E2". Searching the Human Nomenclature Database returns the gene symbol DBT. Searching the Mouse Genome Informatics resources returns the gene symbol Dbt. Therefore, an appropriate symbol adhering to the zebrafish nomenclature guidelines would be dbt.

This is only a starting point for proposing a good symbol. You aren't done yet!

6. If the gene you are naming is a member of a gene family, check to see what is used for the "root" symbol of the gene family in zebrafish. If there is no established gene family in zebrafish but there is in human or mouse, use the same "root" symbol established in human and mouse.
Example: You isolate a new calcium channel, voltage-dependent, alpha 1 subunit gene. Searching ZFIN using the gene query form, you find that all the known calcium channel, voltage-dependent, alpha 1 subunits begin with "cacna1". The last member in the cacna1 series is cacna1f. Therefore, a good starting point would be to propose cacna1g.

7. We've already seen how to propose a zebrafish gene symbol in cases where it is a duplicate, an ortholog of a known human or mouse gene, or when it is a member of a gene family. What if your gene is none of these?
Example: You isolate a novel gene that is an apoptosis determining protein expressed in neurons. You want to call it "apoptosis determining protein, neuronal". adpn would be a starting point.

8. Check ZFIN to ensure that the symbol you want to use is not already in use for a different zebrafish gene or mutant. Also, check the Human Nomenclature Database and the Mouse Genome Informatics resources to ensure the symbol you want to use is not already in use for a different gene in human or mouse.
Example from above: dihydrolipoamide branched chain transacylase E2; dbt. You search ZFIN for a gene already named dbt as well as for a mutant whose abbreviation is dbt. A quick search shows that dbt is not already used for a zebrafish gene or mutant. We know DBT and Dbt are dihydrolipoamide branched chain transacylase E2 in human and mouse. The symbol dbt is still OK.

> **Example from above**: calcium channel, voltage-dependent, alpha 1G subunit; cacna1g. You know cacna1g isn't already used for a zebrafish gene or mutant. However, when you search the Human Nomenclature Database, you find a CACNA1G gene in humans and when you search the Mouse Genome Informatics site, you find Cacna1g. If your zebrafish cacna1 gene is not the ortholog of the human CACNA1G and mouse Cacna1g, you need to propose a different symbol.

Example from above: apoptosis determining protein, neuronal; adpn. You find no zebrafish gene or mutant called adpn. However, when you search the Human Nomenclature Database and the Mouse Genome Informatics resources, you find that ADPN/Adpn is already used for the adiponutrin gene. You need to come up with a different symbol for your apoptosis determining protein gene expressed in neurons.

9. Make sure the symbol you select does not interrupt an existing gene family in zebrafish, human, or mouse by entering a truncated form of your proposed symbol in a gene query form and searching the databases using the "begins with" operator.

Example: You isolate a heart expressed cytokeratin and you want to propose hexc. hexc is not used for zebrafish, human, or mouse genes. However, you search ZFIN, the Human Nomenclature Database and the Mouse Genome Informatics resources for "hex" using the "begins with" operator. You find that in both human and mouse, there are HEXA/Hexa and HEXB/Hexb genes for hexosaminidase A and hexosaminidase B. hexc for your zebrafish gene is not a good symbol because if another human or mouse hexosaminidase gene is found, it will be assigned HEXC/Hexc.

10. Check other resources to ensure that the symbol you select is not a very commonly used, unapproved symbol. Useful resources to check are:

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene

HUGO Gene Nomenclature Committee: http://www.genenames.org/

OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM

PubMed: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed

11. When you have found a suitable symbol, send it to the zebrafish nomenclature committee for confirmation and final approval (nomenclature@zfin.org). We will then assign and reserve this symbol for your novel gene.

Although we typically have a quick turn around time, please don't wait until the last minute to request approval of a gene name and symbol. Often a seemingly simple request is more complicated than anticipated and may require correspondence with other nomenclature committees or independent analyses on our part.

To help us provide quick turn around time for your requests, the following information is greatly appreciated and will be held in strict confidence:

- Provide a sequence accession number if the sequence is public in an on-line sequence database. It is very helpful if you include the amino acid sequence, nucleotide sequence, or both with your request if they are not publicly available.
- If you have mapped the gene on a mapping panel, provide the map position and on which mapping panel the gene is mapped. If you have placed the gene on the Sanger Institute's Ensembl contig (http://www.ensembl.org/Danio_rerio/), please provide the coordinates.
- If you believe the zebrafish gene is the ortholog of a gene in another species, provide the evidence (amino acid sequence comparison, nucleotide sequence comparison, conserved synteny, probe cross hybridization, etc.).
- If possible, please provide a phylogenetic tree or description of the relationships between your zebrafish gene and orthologous or related genes in other species.
- Include any other information or commentary you view as pertinent in assigning proper gene nomenclature.

If you need help with any of this analysis, contact us (nomenclature@zfin.org).