FROG Framework
User Manual
I. Introduction

FROG stands for FingeRprinting Ontology of Genomic variations. FROG fingerprints have been devised to capture all the variations in the genome at all levels i.e. chromosome, DNA, RNA, protein etc. This user manual describes the structure of FROG fingerprints, various features of framework and available search tools designed to search genomics variants based on fingerprints.

The key features of the FROG fingerprints are:

• Organism independent.

• Uses simpler sub-levels vocabularies which can be combined to propose a complex variation effect.

• The sub-levels are placed so that they can include any modifications in future.

• There is also integration of tools like SIFT and POLYPHEN so that the outcome of structural changes can be associated to pathogenicity.

• Many important terms from variation ontology such as VariO and PAGE-OM were carefully curated and placed in hierarchy.

• Genomic variant properties are converted into fingerprints (binary codes) to make it computer efficient in terms of storage and search.
II. Structure of FROG

FROG consists of six major blocks, divided into 48 attributes and 278 properties. The six major blocks are:

1. Chromosome
2. DNA
3. RNA
4. Protein
5. Variation
6. Interaction

Each property of every block is coded by combination of binary codes i.e. 0 and 1. In this way, all the 278 properties are represented by 102 bits called FROG fingerprints. The number of fingerprints used in each block is described in the following table:

<table>
<thead>
<tr>
<th>Block</th>
<th>No. of bits in FROG fingerprints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome</td>
<td>13</td>
</tr>
<tr>
<td>DNA</td>
<td>15</td>
</tr>
<tr>
<td>RNA</td>
<td>23</td>
</tr>
<tr>
<td>Protein</td>
<td>22</td>
</tr>
<tr>
<td>Variation</td>
<td>18</td>
</tr>
<tr>
<td>Interaction</td>
<td>11</td>
</tr>
<tr>
<td>TOTAL</td>
<td>102</td>
</tr>
</tbody>
</table>
III. Block: Chromosome

Chromosome block has five attributes with 38 properties as shown in figure.
IV. Block: DNA

DNA block has seven attributes with 36 properties as shown in figure.
V. Block: RNA

RNA block has nine attributes with 58 properties as shown in figure.
VI. Block: Protein

Protein block has eight attributes with 78 properties as shown in figure.
VII. Block: Variations

Variation block has fifteen attributes with 36 properties as shown in figure

- Reported as model
- Not reported
- Non-genetic origin
- Paternal
- Maternal
- Both
- Tissue details available
- Tissue details not available
- Streched
- Single nucleotide
- Synonymous
- Non-synonymous
- YRI
- JPT
- CHB
- CEPH
- > 1%
- < 1%
- Available
- Not available
- Repeats disease associated pathogenicity association
- Disease causing pathogenicity association
- Associated variants disease associated pathogenicity association
- Not related to clinical phenotype pathogenicity association

Variation block diagram: C. elegans, D. melanogaster, Mus musculus, D. rerio, CPV data, MAF data, Conservation, Details, Geographic details, Origin, Size, Type, Experimental classification, Mesh term.
VIII. Block: Interactions

Interactions block has four attributes with 32 properties as shown in figure.

- Double non-monotonic genetic interaction defined by inequality
- Non-interactive genetic variation defined by inequality - Neutral genetic interaction
- Additive genetic interaction defined by inequality
- Conditional genetic interaction defined by inequality
- Asynthetic genetic interaction defined by inequality
- Single non-monotonic genetic interaction defined by inequality
- Colocalization
- Suppressive genetic variation defined by inequality - Epistatic genetic interaction
- Association
- Epistatic genetic variation defined by inequality - Epistatic genetic interaction
- Synthetic genetic variation defined by inequality - Negative interaction
- Inequality
- Negative interaction
- Neutral interaction
- Not defined
- Positive interaction
IX. FROG web interface

The FROG framework is accessible at http://ab-openlab.csir.res.in/frog and its homepage is divided into four different panels as shown below

1. **Navigation Bar:** It provides links to major components including Search Tools page, User login and Help sections.

2. **Tool panel:** This panel lists the search tools available to the user. These search tools have been divided into two different categories i.e. **Protein variations search** and **RNA variations search.** Throughout this user manual, the names of tools and categories have been written in **bold italics** and **UPPERCASE** respectively.

3. **Detail Panel:** This panel displays the interfaces of search tools for selecting fingerprints. It also shows output of the tool after execution when user clicks on an eye icon shown in history panel.

4. **History Panel:** The history panel shows links to the results of search tools. Each tool generates a new history item which can be used downloaded or visualized in detail panel.
Figure 1: FROG Homepage
X. Creating User Account

- It is recommended that user register their account before using framework. Although unregistered users have access to tools available but their history is stored temporarily. On the other hand, registered users can save and retrieve their results in history panel later too.
- Please note that it is not mandatory to use real email addresses for registration.

Figure 2: User registration form view
XI. Visualization of FROG structure

Frog interface provides a way to explore the structure of FROG fingerprints in an interactive manner as shown in figures below

Figure 3: Collapsed view of FROG structure
Figure 4: Expanded view of FROG structure
In the current version of FROG framework, only mitochondrial variants from MitoLSDB can be searched using fingerprints of different blocks. User selects the different combination of ontology terms which are converted into FROG fingerprints by search tools. Next, these search tools query the database to search for variants with same fingerprints and generate a list of variants as output. The search tools have been divided into two following categories:

- **PROTEIN VARIATIONS SEARCH**: The tools available in this category can be used for searching mitochondrial protein variants using ontology terms of various blocks.
  - **DNA Fingerprints Search**
  - **Protein Fingerprints Search**
  - **Variation Fingerprints Search**

- **RNA VARIATIONS SEARCH**: The tools available in this category can be used for searching mitochondrial RNA variants using ontology terms of various blocks.
  - **DNA Fingerprints Search**
  - **Variation Fingerprints Search**
XIII. DNA Fingerprints Search Tool

As an example, we showcase the use of one of the search tools i.e. DNA Fingerprints Search tool. To find all the mitochondrial protein variants caused by thymine to cytosine transition in DNA, follow the following steps:

1. Click on DNA Fingerprints Search tool under PROTEIN VARIATIONS SEARCH.

2. Keep all the options by default (T --> C is already selected) and click on execute. The tool will report around 21,000 mitochondrial protein variants caused by T to C transition.

Figure 5: DNA Fingerprints Search tool view
Figure 6: Output of DNA Fingerprints Search tool

Similarly, other tools of FROG can be used for searching mitochondrial protein and RNA variants.
XIV. Contact

Questions and suggestions are welcome by FROG Team. For any query or questions, send us an email with full details of tool used and its selected parameters.

FROG Team :

1. Anshu Bhardwaj,
   Scientist,
   CSIR-Open Source Drug Discovery Unit,
   Delhi, India
   Email : anshu@osdd.net

2. E. Abinaya
   Department of BioInformatics,
   Sastra University,
   Tamil Nadu, India
   Email : abinaya.elavarasan@gmail.com

3. Pankaj Narang
   DBT-BINC Fellow,
   School of Computational and Integrative Sciences,
   Jawaharlal Nehru University,
   Delhi, India
   Email : pankaj27_sit@jnu.ac.in