**ACTG 1.10: Novel peptide mapping onto gene models**

**Release: Dec. 19, 2016**

**Hanyang University, Seoul, Korea**

**1. Requirement**

*1.1. Software*

ACTG requires Java Running Environment (JRE) 1.7.0\_45 or later. If JRE is not installed, please visit http://www.oracle.com/technetwork/java/index.html.

*1.2. Hardware*

ACTG’s hardware requirement depends on a size of transcriptome model (GTF). If it is Ensembl transcriptome model, ACTG requires at least 8GB heap size, and 2MB stack size to construct a variant splice graph and to map on the graph.

**2. Usage**

*2.1. Constructing the variant splice graph*

java -Xmx8G -Xss2m -jar ACTG\_Construction.jar const\_params.xml

*2.2. Mapping on the variant splice graph*

java -Xmx8G -Xss2m -jar ACTG\_Mapping.jar mapping\_params.xml

**3. Parameters**

Two parameter files (XML format) are given for construction and mapping separately.

*3.1. const\_params.xml*

|  |  |  |  |
| --- | --- | --- | --- |
| **Tag** | **Type** | **Format** | **Explanation** |
| Input | Transcriptome | GTF | Path of a folder which contains transcriptome models. |
| Input | Reference  Genome | FASTA | Path of a folder which contains reference genome. |
| Output | Graph file | Ser | Name of a constructed variant splice graph file. |
| TheNumber  OfThreads | Setting the number of threads. A proper setting can increase the speed of construction. | | |

*3.2. mapping\_params.xml*

|  |  |  |  |
| --- | --- | --- | --- |
| **Tag** | **Type** | **Format** | **Explanation** |
| MappingMethod | - | - | PV: Mapping [P]rotein database first, then next [V]ariant splice graph.  PS: Mapping [P]rotein database first, then next [S]ix-frame translation.  VO: Mapping [V]ariant splice graph [O]nly.  SO: Mapping [S]ix-frame translation [O]nly. |
| ILSame | - | - | Yes/No: consider isoleucine and leucine as the same. |
| Input | Peptide list | List | Path of a peptide list file that a user wants to map. |
| Output | outputPath | - | Path of a folder where the output files will be located. |
| Input | Protein  Database | FASTA | If a user sets “MappingMethod” as P[V|S], then the user should provide a folder path containing protein database. |
| SAVs | - | - | Yes/No: automatically consider single amino-acid variant when mapping on protein database. |
| Input | Graph file | Ser | Serialization file path of the variant splice graph. |
| Junction Variation | - | - | Yes/No: consider junction variation events. |
| Exon Skipping | - | - | Yes/No: consider single exon skipping events. |
| Intron Mapping | - | - | Yes/No: consider intron mapping (exon-extension) events. |
| Mutation | - | - | If a user gives a VCF (variant call format), then this setting must be “yes”. |
| Input | Mutation | VCF | If “Mutation” tag is set as “yes”, then a user should provide the path of folder where VCF files are. |
| Input | Reference Genome | FASTA | If a user sets “MappingMethod” as PS or SO, then the user should provide the path of folder where reference genome files are. |
| TheNumber  OfThreads | A user can set the number of threads. A proper setting can increase the speed of mapping. | | |