VarDict Getting Started Guide

Version Number: 1.0

|  |  |
| --- | --- |
| Document Details | |
| Document Title | VarDict Getting Started Guide |
| Status / Version | Final/1.0 |
| Reference Code |  |

|  |
| --- |
| Disclaimer |
| If you are reading a copy of this document (i.e. print out of controlled electronic document or an electronic copy of a wet ink signed paper document) you may be reading an uncontrolled version. Before making updates, please ensure that you are using the latest version of this document and the associated template. |

|  |  |  |  |
| --- | --- | --- | --- |
| Document Approval / Review | | | |
| Criteria | Name & Role | Signature | Date |
| Author | Viktor Kirst |  | 11-02-2015 |
| Approved |  |  |  |
| Approved |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Version Control | | | |
| Version | Date | Description of Changes | Author |
| 0.1 | 11/02/15 | Initial version | Viktor Kirst |
| 1.0 | 12/02/15 | Review and formatting | Liudmila Simonova |
|  |  |  |  |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Referenced Documents | | | |
| No. | Document Title and ID | Version | Location |
|  |  |  |  |
|  |  |  |  |

**Table of Contents**

**TOC \o 2-2 \t "Heading, 3"**

1. **Source Location PAGEREF \_Toc \h 3**
2. **Build Instruction PAGEREF \_Toc1 \h 3**
   1. **Distribution Package Structure PAGEREF \_Toc2 \h 3**
   2. **Third-Party Libraries PAGEREF \_Toc3 \h 3**
3. **Launch Instruction PAGEREF \_Toc4 \h 3**

# Source Location

The VarDict source code is located at <https://github.com/AstraZeneca-NGS/VarDictJava>.

To load the project, execute the following command:

git clone —recursive https://github.com/AstraZeneca-NGS/VarDictJava.git

Note that original VardDict project is placed in this repository as submodule and it's contents can be found in sub-directory VarDict in VarDictJava working folder. So when you use teststrandbias.R and var2vcf\_valid.pl. (see details and examples below), you have to add prefix VarDict: VarDict/teststrandbias.R and VarDict/var2vcf\_valid.pl.

# Build Instruction

The following prerequisites are necessary to build the VarDict project:

* JDK1.7 (or higher)
* Internet access

The project uses Gradle (<http://gradle.org/>) and already includes a gradlew script.

To build the project, in the root folder of the project, run the following command:  
./gradlew clean installApp.

## Distribution Package Structure

When the build command completes successfully, the build/install/VarDict folder contains the distribution package.

The distribution package has the following structure:

bin/ - contains the launch scripts

lib/ - has the jar file that contains the compiled project code and the jar files of the third-party libraries that the project uses.

You can move the distribution package (the content of the build/install/VarDict folder) to any convenient location.

## Third-Party Libraries

Currently, the project uses the following third-party libraries:

* **JRegex** (<http://jregex.sourceforge.net>, BSD license) is a regular expressions library that is used instead of the standard Java library because its performance is much higher than that of the standard library.
* **Commons CLI** (<http://commons.apache.org/proper/commons-cli>, Apache License) – a library for parsing the command line.

# Launch Instruction

The following prerequisites are necessary to launch the program:

* Java 1.7 (or higher)
* Samtools (must be in path)

To launch VarDict:

* Run the command <distribution package path>/bin/VarDict or
* Run just VarDict if <distribution package path>/bin/ is in path

The work of the program is identical to the function of the original Perl script and thus can replace it.

For example, the original launch command (from the ***readme*** file of the perl version) is:

vardict -G /path/to/hg19.fa -f 0.01 -N sample\_name -b /path/to/my.bam -z -c 1 -S 2 -E 3 -g 4 /path/to/my.bed | teststrandbias.R | var2vcf\_valid.pl -N sample\_name -E -f 0.01

The same command in VarDictJava looks like:

<distribution package path>/bin/VarDict -G /path/to/hg19.fa -f 0.01 -N sample\_name -b /path/to/my.bam -z -c 1 -S 2 -E 3 -g 4 /path/to/my.bed | VarDict/teststrandbias.R | VarDict/var2vcf\_valid.pl -N sample\_name -E -f 0.01

The program uses the same parameters as the Perl version. The Java version contains an additional parameter: -th (from ‘threads’). This parameter manages the number of threads that do the work. If this parameter is missing, the number of threads equals to the number of processor cores.