Association analysis and more with TASSEL

M. Cinta Romay
05/10/2012
TASSEL:
1. What is Tassel?
2. Where do I get Tassel?
3. How do I use Tassel?
   a. Graphical User Interface (GUI)
   b. Standalone version
   c. Source code
1. What is Tassel?

- Java application that runs in your computer (any operating system)
- Trait Analysis by aSSociation, Evolution and Linkage: provides tools to investigate the relationships between genotypes and phenotypes
  a. Association study
  b. Analysis of Linkage Disequilibrium
  c. Principal Component Analysis
  d. Cluster Analysis
  e. Overview of genotypic data
  f. ...
- Design to analyze large datasets
2. Where do I get Tassel?

www.maizegenetics.net/tassel
2. Where do I get Tassel?

www.maizegenetics.net/tassel
2. Where do I get Tassel?

www.maizegenetics.net/tassel
2. Where do I get Tassel?

www.maizegenetics.net/tassel
2. Where do I get Tassel?

www.maizegenetics.net/tassel

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 3.0 (Updated: December 22, 2011)
Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation
Tassel 3.0 Pipeline Document (Updated April 26, 2012)
Tassel 3.0 GES Pipeline Document (Updated March 22, 2012)
Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)
Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipline.jar (Updated May 19, 2010)

Tassel 3.0 Video Tutorials
Demonstration of MLM Association Analysis with Gramene Diversity Data
2. Where do I get Tassel?

www.maizegenetics.net/tassel

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 3.0 (Updated: December 22, 2011)

Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation

Tassel 3.0 Pipeline Document (Updated April 26, 2012)
Tassel 3.0 GES Pipeline Document (Updated March 22, 2012)
Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)
Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipeline.jar (Updated May 19, 2010)

Tassel 3.0 Video Tutorials

Demonstration of MLM Association Analysis with Gramene Diversity Data
2. Where do I get Tassel?

www.maizegenetics.net/tassel

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 2.0 (Updated: December 22, 2011)

Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation

Tassel 3.0 Pipeline Document (Updated April 26, 2012)
Tassel 3.0 GBS Pipeline Document (Updated March 22, 2012)
Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)
Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipeline.jar (Updated May 19, 2010)

Tassel 3.0 Video Tutorials
Demonstration of MLM Association Analysis with Gramene Diversity Data
2. Where do I get Tassel?

www.maizegenetics.net/tassel

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 3.0 (Updated: December 22, 2011)
Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation
Tassel 3.0 Pipeline Document (Updated April 26, 2012)
Tassel 3.0 GES Pipeline Document (Updated March 22, 2012)
Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)

Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipeline.jar (Updated May 19, 2010)

Tassel 3.0 Video Tutorials
Demonstration of MLM Association Analysis with Gramene Diversity Data
2. Where do I get Tassel?

www.maizegenetics.net/tassel

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 3.0 (Updated: December 22, 2011)
Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation
Tassel 3.0 Pipeline Document (Updated April 26, 2012)
Tassel 3.0 GES Pipeline Document (Updated March 22, 2012)
Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)
Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipeline.jar (Updated May 19, 2010)

Tassel 3.0 Video Tutorials
Demonstration of MLM Association Analysis with Gramene Diversity Data
2. Where do I get Tassel?

www.maizegenetics.net/tassel

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 2.0 (Updated: December 22, 2011)
Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation
Tassel 3.0 Pipeline Document (Updated April 26, 2012)
Tassel 3.0 GES Pipeline Document (Updated March 22, 2012)
Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)
Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipeline.jar (Updated May 19, 2010)

Tassel 3.0 Video Tutorials
Demonstration of MLM Association Analysis with Gramene Diversity Data
2. Where do I get Tassel?

www.maizegenetics.net/tassel

Contacts

We recommend posting questions to the discussion group

Tassel User Group (tassel@googlegroups.com)

General Questions / Pipeline
Terry Casstevens (tmc46@cornell.edu)

Analysis Tools / MLM / GLM
Peter Bradbury (pjb39@cornell.edu)
Zhiwu Zhang (zz19@cornell.edu)

Overall Project Lead
Ed Buckler (esb33@cornell.edu)
3. How do I use tassel?

www.maizegenetics.net/tassel
3. How do I use Tassel?

www.maizegenetics.net/tassel
3. How do I use tassel?

www.maizegenetics.net/tassel

Do not panic!!

Some incompatibilities might not allow your java virtual machine to start
3. How do I use tassel?

www.maizegenetics.net/tassel
3. How do I use tassel?

www.maizegenetics.net/tassel
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use Tassel?
   a. Graphical User Interface (GUI)

   - Datasets and reports - Stored in RAM
   - Task progress
   - Comments about data and reports
3. How do I use tassel?
   a. Graphical User Interface (GUI)

Functions are organized in 3 modes
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use Tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)

   ![Filter Alignment Interface]

   - Minimum Count: 1
   - Minimum Frequency: 0.0
   - Position Type: Position index
   - Start Position: 0
   - End Position: 43960
   - 5% MAF: 0.05
   - Total sequences: 2815
   - Total sites: 43960

   [Filter] [Cancel]
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?

a. **Graphical User Interface (GUI)**
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)

TASSEL v 4.0
3. How do I use tassel?

a. Graphical User Interface (GUI)

TASSEL v 4.0
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?

a. Graphical User Interface (GUI)

TASSEL v 4.0
3. How do I use tassel?

a. Graphical User Interface (GUI)

TASSEL v 4.0
3. How do I use tassel?

a. Graphical User Interface (GUI)

TASSEL v 4.0
3. How do I use tassel?

a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)
3. How do I use tassel?
   a. Graphical User Interface (GUI)

Steps to perform Association analysis:

1. Load genotypic data
2. Load phenotypic data
3. Load population structure data or run a PC analysis
4. Merge G+P+Q
5. Run GLM (simple trait)
6. Load Kinship matrix/ calculate kinship matrix
7. Select G+P+Q and K and run MLM
8. Visualize the results with QQ plot and Manhattan plot
3. How do I use tassel?

b. Standalone version

- Consists on modules (i.e. Plugins)
- Output from one module can be input to another module. Determined by order specified.

```
run_pipeline.bat -fork1 -h input.txt -ld -ldd png -o output.png -runfork1
```

GBS analysis pipeline is an extension to Tassel, and, as such, GBS commands are run as TASSEL plugins via the command line

* Terry Casstevens (tmc46@cornell.edu)
3. How do I use tassel?
   b. Standalone version

   • Install:
     • Unzip the Tassel standalone build onto your system
   • Change into the root directory:
     • `tasselx.0.standalone`
   • Execute:
     • Windows: use `run_pipeline.bat`
     • UNIX: `run_pipeline.pl`

* Terry Casstevens (tmc46@cornell.edu)
3. How do I use tassel?

b. Standalone version

- Modify memory from bat file:

```bash
@echo off
set TOP=
set LIB_JARS=%TOP%\lib
set CP=%TOP%\sTASSEL.jar
for %i in (%LIB_JARS%\*.jar) do call "%TOP%\cp.bat" %i
echo %CP%
java -classpath "%CP%" -Xms512m -Xmx1536m net.maizegenetics.pipeline.TasselPipeline %*
```

Similarly for the perl file: edit and find the –Xmx command!

* Terry Casstevens (tmc46@cornell.edu)
3. How do I use tassel?

b. Standalone version

- Pipelines consist of pipeline segments. Each segment has a name (1, A...) that follows:
  - -fork: start of a pipeline segment that should be executed sequentially
  - -combine: starts a pipeline segment that combines datasets from multiple pipeline segments
- Output from a pipeline segment can be used as input to another by referencing with -input

```
run_pipeline.bat –fork1 ...–fork2... -fork3... –combineA –input1 –input2 ...
```

* Terry Casstevens (tmc46@cornell.edu)
3. How do I use tassel?

b. Standalone version

- Pipeline segments can initiate execution in two ways:
  - `runforkX`
  - By receiving input from another segment

```bash
run_pipeline.bat -fork1 -h "path/File1.txt" -fork2 -includeTaxaInFile "path/TaxaFile1.txt" -input1 -export "path/OutputFile1.txt" -exportType Hapmap -fork3 -includeTaxaInFile "path/TaxaFile2.txt" -input1 -export "path/OutputFile2.txt" -exportType Hapmap -runfork1
```

* Terry Casstevens (tmc46@cornell.edu)
3. How do I use tassel?
   b. Standalone version

To Report Bugs or Request Features, please go to sourceforge, and click "Add new"...

Tassel Bug Tracking
Tassel Feature Requests

Tassel Documentation 3.0 (Updated: December 22, 2011)
Tassel User's Guide
Tassel Tutorial Data
Visit Tassel User Group
Tassel FAQs
Tassel JAVA API (Updated April 26, 2012)

Tassel Pipeline 3.0 Documentation
   Tassel 3.0 Pipeline Document (Updated April 26, 2012)
   Tassel 3.0 GBS Pipeline Document (Updated March 22, 2012)
   Tassel 3.0 UNEAK Pipeline Document (Updated April 12, 2012)
   Tassel 3.0 Pipeline Tutorial (Updated May 11, 2011)
   Tassel 3.0 Creating Self-Describing Plugins (Updated June 2, 2011)

Tassel Import/Export Command Line Interface. For Usage: java -jar AlleleFileConversionPipeline.jar (Updated May 19, 2010)
3. How do I use tassel?

c. Source code

- Allows manipulation of the code to adapt it to your own projects (requires some java/coding knowledge)

- Sourcecode is publicly available:
  - [https://tassel.svn.sourceforge.net/svnroot/tassel](https://tassel.svn.sourceforge.net/svnroot/tassel)

- Instructions for setting up Tassel in NetBeans can be found at:
  - [http://www.maizegenetics.net/tassel-in-netbeans](http://www.maizegenetics.net/tassel-in-netbeans)
3. How do I use tassel?

c. Source code

```java
Pattern splitter = Pattern.compile("; ");
int numTaxa = alignment.getSequenceCount();
for (int taxa = 0; taxa < numTaxa; taxa++) {
    String[] name = splitter.split(alignment.getAlignment(taxa), getFullName()).trim();
    // if name.length > 1 { // name1 is readable
    PEDbw.write(name[1]); // name1 is readable
    }
    else {
        PEDbw.write("-9");
    }
    PEDbw.write(delimChar);
    // PEDbw.write(name[0]); // name1 is readable
    PEDbw.write(alignment.getAlignment(taxa), getFullName()).trim(); // name1 is readable
    PEDbw.write(delimChar);
    // PEDbw.write("-9"); // paternal ID available
    PEDbw.write(delimChar);
    PEDbw.write("-9"); // maternal ID available
    PEDbw.write(delimChar);
    // gender is both
    PEDbw.write(delimChar);
    PEDbw.write("-9"); // phenotype unavailable, changed to "-9" from "NA" due to error with MapMain
    PEDbw.write(delimChar);
    for (int site = 0; site < numSites; site++) {
        byte[] b = AllelePositionBLOBUtils.getSNPValueForPlink(AllelePositionBLOBUtils.getSNPValueForPlink()
            PEDbw.write((char)b[0]);
            PEDbw.write(delimChar);
            PEDbw.write((char)b[b.length - 1]);
            if (site == numSites - 1) {
                PEDbw.write(delimChar);
            }
        }
```