

Développement d'une application web interactive sous R avec le package 'Shiny' pour la comparaison des résultats NGS (Next Generation Sequencing) d'IMGT/HighV-QUEST des récepteurs d'antigènes

Safa Aouinti ^{a,b,c}, Dhafer Malouche ^{b,c} and Marie-Paule Lefranc ^a

^aIMGT[®], the international ImMunoGeneTics information system[®],
Institut de Génétique Humaine, UPR CNRS 1142, Université Montpellier 2
141 rue de la Cardonille 34396 Montpellier Cedex 5- France
safa.aouinti@igh.cnrs.fr, marie-paule.lefranc@igh.cnrs.fr

^bDépartement Statistique
École Supérieure de la Statistique et de l'Analyse de l'Information de Tunis, Université de Carthage
6, Rue des Métiers- Charguia II- B.P 675, 1080 Tunis- Tunisie
dhafer.malouche@me.com

^cDépartement des Technologies de l'Information et de la Communication
École Nationale d'Ingénieurs de Tunis- Unité de recherche Signaux et Systèmes
BP 37, Le Bélvédère 1002 Tunis- Tunisie

Mots clefs : Statistique, package Shiny, test de comparaison de deux proportions, intervalle de confiance pour une différence de deux proportions, visualisation, IMGT/HighV-QUEST.

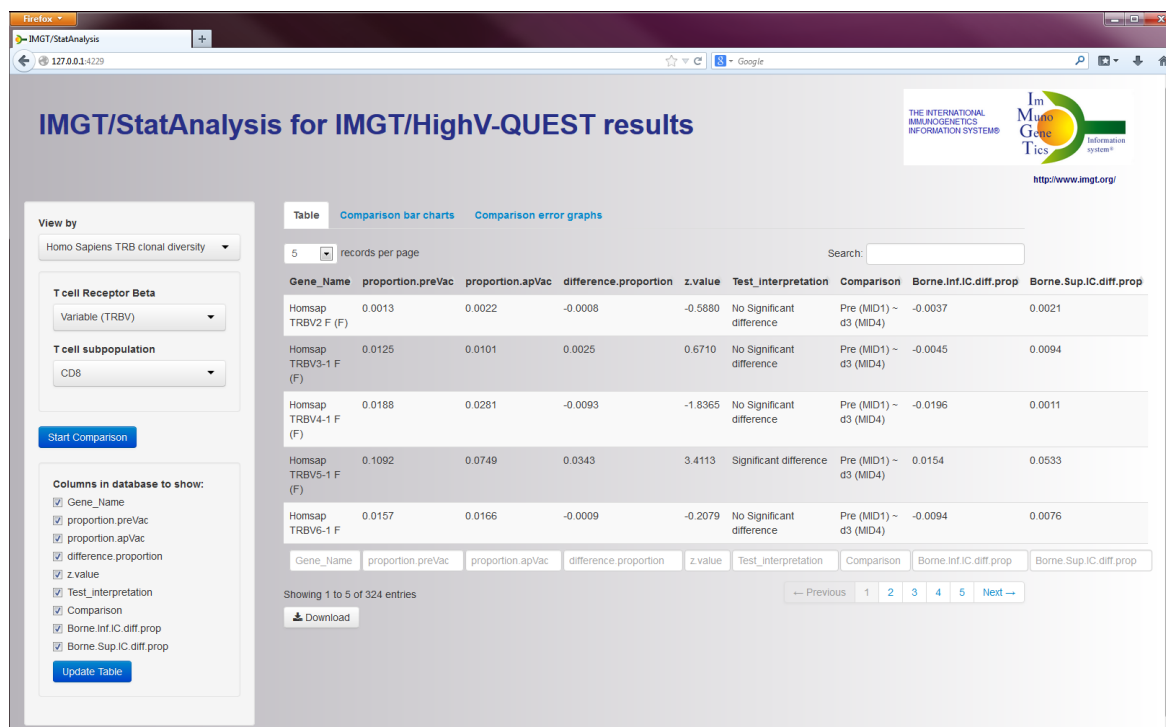
Le package Shiny est un package libre développé par RStudio (<http://www.rstudio.com/>) qui rend le logiciel R facile d'utilisation pour des requêtes et des résumés de données interactifs, par le biais d'un navigateur web. Il est livré avec une variété de widgets pour créer rapidement des interfaces interactives (<http://shiny.rstudio.com/>). Le style par défaut d'une application Shiny est efficace, mais Shiny est très extensible et il est facile pour l'utilisateur d'intégrer des applications avec son propre contenu web en utilisant les langages HTML et CSS. JavaScript et jQuery peuvent être utilisés pour étendre la portée des applications. Les utilisateurs de ces applications choisissent les paramètres d'entrées à partir de commandes faciles tels que les curseurs, les listes déroulantes et les champs de texte [1][2].

Le but de ce *lightning talk* est de présenter une application développée avec le package Shiny pour visualiser les résultats issus d'une comparaison de sets de séquences NGS (*Next Generation Sequencing*) analysées par IMGT/HighV-QUEST.

IMGT/HighV-QUEST est le portail de référence pour l'analyse haut débit des séquences d'immunoglobulines (IG) et récepteurs T (TR) obtenues par les technologies de séquençage NGS. Cette approche permet d'appréhender pour la première fois la complexité et la diversité de la réponse immunitaire qui est quasiment illimitée (2.10^{12} IG et TR différents par individu). IMGT/HighV-QUEST analyse des séquences des IG et TR et identifie les gènes variable (V), de diversité (D) et de jonction (J) des domaines spécifiques qui se lient à l'antigène et constituent ainsi le répertoire immunitaire. Ces répertoires contribuent à la comparaison des profils immunitaires en situation normale (vaccination) et pathologique (leucémies, lymphomes, myélomes, maladies autoimmunes, cancer, infections, etc.) [3][4]. IMGT/HighV-QUEST est librement disponible pour les universitaires sur la page d'accueil IMGT[®] (<http://www.imgt.org>) depuis octobre 2010.

Au cours de cette étude, 12 sets de séquences ont été comparés provenant de 3 sous-populations de cellules T à 4 points temporels (avant vaccination H1N1, jour 3, jour 8 et jour 26 après vaccination) d'un même individu [5].

Nous avons développé une interface interactive (Figure) qui permet de visualiser les résultats de la comparaison des 12 sets (deux par deux) suite à l'application du test de comparaison de deux proportions (seuil de signification statistique fixé à $\alpha=5\%$). Dans un premier temps, la comparaison des proportions de séquences entre deux mêmes gènes dans deux sets différents a été visualisée par des diagrammes en barres juxtaposées. Dans un second temps, la différence de proportions entre ces gènes a été représentée par des barres d'erreurs associées aux intervalles de confiance à 95%. Ces dernières ont été tracées pour résumer l'information contenue dans la table des résultats, et ceci, en fonction de la significativité de la différence jugée par le test appliqué. Tous les graphes ont été générés en utilisant le package ggplot2 qui fournit un modèle puissant de graphiques.



Références

- [1] Beeley, C. (2013) Web Application Development with R using Shiny. *Packt Publishing Ltd*-110 pages.
- [2] Paulsonn J., and Cheng, J., Tutoriel : Developing web applications with R and shiny.
- [3] Alamyar, E., Duroux, P., Lefranc, M.-P. and Giudicelli, V. (2012). IMGT® tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations : IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. In : B. Tait and F. Christiansen (Eds.), *Immunogenetics, Humana Press, Springer, New York, USA. Methods Mol. Biol*, chap 32.
- [4] Alamyar, E., Giudicelli, V., Li, S., Duroux, P., Lefranc, M.-P. (2012). IMGT/HighV-QUEST : The IMGT® web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. *Immunome Res.* 8(1) :26.
- [5] Li, S., Lefranc, M.-P., Miles, J.J., & al. (2013).IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nature Communications*, 4 :2333 doi : 10.1038/ncomms3333.