

A bioinformatics workflow for detecting signatures of selection in genomic data

Murray Cadzow^{1,2}, James Boocock^{1,2}, Hoang Tan Nguyen^{1,2},
Phillip Wilcox^{2,3}, Tony R Merriman¹ and Michael A Black¹

¹Department of Biochemistry, University of Otago

²Department of Mathematics and Statistics, University of Otago

³Scion Research, Rotorua, New Zealand

September 9, 2013

Contents

1 Introduction

Mik to fill this in

2 Getting Started

2.1 Installation

To install the package standalone will require manual configuration of the config file run this as root.

```
./install.sh --standalone
```

The rest of this section will be dedicated to the automatic installation. To perform an automatic installation of the selection pipeline run as root.

```
./install.sh
```

Installation creates a default config file located in the base directory of the pipeline. Installation adds a program to the system path selection_pipeline.

To see the help and test the program is installed correctly run the following command.

```
selection_pipeline -h
```

2.2 Selection Signatures at the Lactase Locus

2.2.1 Getting the Data

To perform this tutorial you will be required to install tabix, this package includes the tabix software to install the software on ubuntu requires running the command.

```
apt-get install tabix
```

On other distributions just run as root.

```
./install_tabix.sh
```

This places the tabix executable in the *bin/* folder. to run tabix type *bin/./tabix* in a terminal from the base directory of the selection pipeline.

The lactase gene is located on Chromosome 2 between 136,545,410-136,594,750 positions. For this example we will use a 10 megabase region containing Lactase and the CEU and YRI populations from the 1000 genomes. Usually iHS would be done chromosome wide but in order to demonstrate how to use the pipeline we will use a small part of chromosome 2 between 130-140 megabases. To download this file from 1000 genomes use the command below you will need to use tabix to download this region.