

ADMIXFRQ

Version 1 user manual

Authors:

Rafael A. Nafikov

Ellen M. Wijsman

Division of Medical Genetics

Department of Medicine

University of Washington

In progress!!! Last modified on: March 20, 2018

Contents

1. Introduction	1
1.1 Citing ADMIXFRQ	2
1.2 License	3
1.3 Software URL	4
2. Running ADMIXFRQ	5
2.1 Command line to run ADMIXFRQ	6
2.2 Parameter file	6
2.3 Data input files	6
2.4 Output files.....	8

1. [Introduction](#)

2.1 Command line to run ADMIXFRQ

The ADMIXFRQ script can be run from a terminal using either:

perl ADMIXFRQ.pl parameter_file.txt

or

./ADMIXFRQ.pl parameter_file.txt

command line.

2.2 Parameter file

Parameter file is a space or tab delimited text file which contains necessary information for the script to perform computations. Curly brackets { } are only used to highlight certain symbols or entries used in a parameter file and must not to be used in a parameter file. Empty lines or lines starting with ampersand { # } are ignored. Any text entered after ampersand { # } on a line in a parameter file is ignored. Each line of a parameter file contains either one or multiple options. Each option has its corresponding entry separated by either equal sign { = } or colon { : }. Multiple options entered on a single line of a parameter file are separated by a vertical line { | }. Arrangement of multiple options on a single line was dictated by their interrelatedness.

Required options:

➤ **CHROMOSOMES = [1 2 3 7 9]**

Enter individual chromosome numbers inside square brackets separated by space. Alternatively, consecutive chromosome numbers can be specified inside square brackets as a range of numbers using two dots .. and no spaces between the dots and numbers (ex. **[1..5]** means chromosomes 1, 2, 3, 4, and 5). It is also allowed to use the above two specification formats for chromosome number simultaneously (ex. **[1..5 9 13 17]**).

➤ **FAMILY NAMES = [family_name1 family_name2]**

Enter family names inside square brackets separated by space. Exactly the same family names must be used in parameter and data files.

➤ **ANCESTRY MODEL = Global FBCW FBGW Local | AVERAGE TYPE FOR GLOBAL AND FBGW = normal**

- ❖ Enter admixture (ancestry) model names, you would like to use in the analysis, separated by space. The available options for admixture (ancestry) models are: **Global FBCW FBGW Local** . A definition of each of the models can be found in Nafikov et. al (2018). Options for different admixture models were provided for exploratory purposes. For data analysis we recommend to use only Local admixture model because, in our opinion, it provides the most appropriate results for the analysis.

➤ **EXTRAPOLATION OF LOCAL ADMIXTURE PROPORTIONS: yes**

There are only two options for this entry: **yes** or **no** . If **yes** option is used and you do not have admixture estimates covering a full range of your marker data for which you would like to modify allele frequencies and calculate admixture proportions, then admixture proportions estimated at your first and last marker positions will be extrapolated to the beginning and end of a chromosome

➤ **GLOBAL ANCESTRY MODEL PROPORTIONS: EUR = 0.64 | AFR = 0.27 | AMR = 0.09**

This option can be used to specify admixture proportions for reference populations assumed to be involved in admixture under the Global ancestry (admixture) model. The specified admixture proportions will be used in calculation of marker allele frequencies under the Global ancestry (admixture) model.

➤ **COMPUTE GLOBAL ANCESTRY MODEL PROPORTIONS: yes**

This option can be used to compute admixture proportions under the Global ancestry (admixture) model. This computation requires genome-wide local ancestry estimates for all the families under consideration.

➤ **GENO FILE FOR GL_AUTO: FILE PATH = /your_directory/+/chr*/chr*.geno | DETAILED ALLELE FREQUENCY FILE = yes**

➤ **ADMIXTURE PROPORTIONS PRINTOUT: yes**

- **INFORMATION FOR GENO FILE: FILE PATH =**
/your_directory/chr*/chr*.tmap | HEADER = F | BP_POSITION = 4 |
CM_POSITION = 3 | RS_NUMBER = 2

- **MODIFY ALLELE FREQUENCIES: FILE PATH = /your_directory/chr*.txt |**
HEADER = T | BP_POSITION = 2

- **OUTPUT DIRECTORY = /your_output_directory**

- **YOUR SAMPLE ALLELE FREQUENCIES: FILE PATH =**
/your_directory/chr*.txt | HEADER = T | BP_POSITION = 2 | ALELLE 1 = 5 |
FREQUENCY OF ALLELE 1 = 6 | ALELLE 2 = 7 | FREQUENCY OF ALLELE
2 = 8

- **SUBPOPULATION NAME = EUR | FILE PATH = /your_directory/chr*.txt |**
HEADER = T | BP_POSITION = 3 | RS_NUMBER = 2 | REF_ALLELE = 4 |
ALT_ALLELE = 6 | REF_ALLELE_FRQ = 5 | ALT_ALLELE_FRQ = 7

- **LOCAL ADMIXTURE PER FAMILY FOR SUBPOPULATION = EUR | FILE**
PATH = /your_directory/chr*.txt | HEADER = T | VARIANT IDENTIFIER =
rs_number | VARIANT IDENTIFIER COLUMN = 1 | FAMILIES = [CU0007F
CU0044F CU0068F] | COLUMN NUMBERS = [2 3 4]

- **EXTRA INFORMATION FOR ADMIXTURE FILES: FILE PATH =**
/your_directory/extra_information.txt | HEADER = T | RS_NUMBER = 2 |
BP_POSITION = 4 | CHR_NUMBER COLUMN = 1

