

---

# EMIBD9 (Linux)

---

## Introduction

This document briefly describes how to run EMIBD9 on a Linux machine. It should be read together with the EMIBD9UserGuide.pdf file, which gives much more details about the methods, inputs and outputs of EMIBD9 and is thus highly recommended to read before starting the use of EMIBD9.

The zipped Linux package includes the following files in 3 subfolders:

- 1) EM\_IBD\_P
- 2) EM\_IBD\_P\_impi
- 3) EM\_IBD\_P\_openmpi
- 4) Readme.pdf
- 5) EMIBD9UserGuide.pdf
- 6) EMIBD9UpdateHistory.txt
- 7) ant377.dat
- 8) ant377.par

The first 3 files are the binaries compiled by Intel's Fortran compiler (2018 Update 3). They are all for 64bit Linux, and may not work for 32bit Linux. The 1<sup>st</sup> file is statically linked for serial run, the 2<sup>nd</sup> file is linked with Intel's MPI library (2018 Update 3) and the 3<sup>rd</sup> file is linked with openMPI library (3.1.4). The latter 2 files are for MPI parallel run, using the corresponding MPI libraries installed on your cluster. The 3 binary files are in the subfolder "Bin".

The 4th file is what you are reading. The 5th is a detailed description of the use of EMIBD9 program. The 6th file is a brief updating history of EMIBD9. These 3 documents are in the subfolder "Doc".

The 7th file is an example genotype data file, and the 8<sup>th</sup> is the corresponding parameter file. These 2 files are in the subfolder "Example".

## Serial run

To make a serial run of EMIBD9 for the example dataset in subfolder Example, simply open an x terminal and navigate to the Example folder (using command cd), and type

```
*/EM_IBD_P INP:ant377.par
```

where "\*" is the path of the EMIBD9 binary files. For example, the path could be "\$HOME/Scratch/EMIBD9/Bin", and the entire command line is thus "\$HOME/Scratch/EMIBD9/Bin/EM\_IBD\_P INP:ant377.par".

If, on installing EMIBD9 program on your computer, you have included the **EMIBD9** program path permanently in the automatic search paths of your shell (as detailed in EMIBD9UserGuide.pdf), then the command line for running the example data simplifies to

```
EM_IBD_P INP:ant377.par
```

On finishing the run, analysis results will be in output file ant377.ibd9 in Example folder.

In the above command line, the key word “INP:” means INPut file. Other key words are listed and explained in the main documentation, EMIBD9UserGuide.pdf.

To run your own dataset, you need to make a project folder, to prepare a parameter input file and a genotype data file, and to save the 2 files in the project folder. Suppose your project folder is “MyPrj”, and in this folder you have a parameter file MyPrj.par and a data file MyPrj.dat. Note the data file name “MyPrj.dat” must be specified in parameter file “MyPrj.par”. You can run the analysis of your data from your project folder, MyPrj, by the X-terminal command line

```
*/EM_IBD_P INP:MyPrj.par
```

On finishing the run, analysis results will be in an output file stored in your project folder, MyPrj.

### **MPI parallel run with multiple processes**

For parallel run with multiple processes, you need either Intel’s MPI (2018, Update 3) or openMPI (3.1.4) installed on your Linux machine. Most often, openMPI is already installed. If not, you need to download (openMPI is free) and install it.

There are several different implementations of MPI. Unfortunately, they are not compatible. A binary produced with one MPI implementation may not work with another MPI implementation. Even with the same implementation (say, openMPI), different versions (e.g. 1.8.4 vs 3.1.4) may also be incompatible. For this reason, I have provided two parallel versions of EMIBD9, compiled and linked to different MPI implementations.

Choosing the right binary suitable for the MPI implementation and version number on your machine, you can start running the example dataset with 4 processes, from Example folder, by typing:

```
mpirun -np 4 */EM_IBD_P_mpi INP:ant377.par
```

or

```
mpirun -np 4 */EM_IBD_P_openmpi INP:ant377.par
```

If you want to use 2 openMP threads per MPI process (4 in total), the command line is

```
mpirun -np 4 */EM_IBD_P_mpi INP:ant377.par OMP:2
```

Note openMP parallel threads can be used with or without using MPI. In the above command line example, a total number of 8 parallel threads will be used with 2 threads with shared memory running within each of the 4 MPI processes.

When executing the serial or MPI parallel binaries in the first time, you may see an error message “permission denied”. This means the binary needs to be unlocked for execution by issuing the X-terminal command line

```
chmod u+x ./Bin/*
```

in the EMIBD9 program folder, where \* is the binary name.