

Manuál pro program

PedOrder

verze : 1.00

Program pro vytvoření čísla generací předků
v rodokmenu pro návazná
genetická vyhodnocování

2018

Autoři: Vostrý L., Příbyl J., Krupa E., Vostrá-Vydrová H.

Výzkumný ústav živočišné výroby (Institute of Animal Science)

Přátelství 815

104 00 Praha – Uhřetěves

Česká republika

Vypracováno za podpory projektu MZe - Celostátní informační systém genetického hodnocení
hospodářských zvířat (QJ1510139)

Úvod

Program PedOrder byl vyvinut za účelem vytvoření označení generací jedinců v souboru rodokmenu, který slouží k seřazení souboru rodokmenu, které je využíváno v mnoha počítačových programech, například v počítačových programech pro odhad populační rozmanitosti, pro sestavování matic vzájemných příbuzností, nebo pro odhad genetických parametrů a v některých programech pro stanovení plemenné hodnoty – DMU (Madsen a Jansen, 2008). Program PedOrder je však univerzálním programem, který umožňuje seřazení rodokmenu dle generací, které může být použito pro mnoho rozdílných přístupů.

Systémové prostředí

Program PedOrder je dostupný ve formě zdrojového kódu programovacího jazyku FORTRAN 90. Zdrojový kód je možné zkompileovat běžným kompilátorem pod licencí GNU (gfortran) pod systémy Linux a Windows (32 a 64 bit) viz kapitola instalace.

Postup instalace

Pro získání zdrojových kódů v balíku pedored.tar.gz je nutné balík se zdrojovými kódy a ukázkovými soubory rozbalit. Ukázkový postup je uveden v prostředí LINUX (UNIX) a pomocí GNU Fortran kompilátoru GNU (gfortran), ale obdobný způsob je možný i pro prostředí WINDOWS a jiné Fortran kompilery (např. ifort aj.).

Rozbalení se provede příkazem

```
tar -xvzf pedorder.tar.gz
```

Po rozbalení balíku je nutné program zkompileovat. Kompilaci provedeme příkazy:

```
cd ./lib/PedOrder
```

```
gfortran pedOrder.f90 -o pedOrder
```

Po zkompileování programu vznikne spustitelný program: **pedOrder**. Tento program je již možné zkopírovat do adresáře s rodokmeny, či do adresáře */usr/local/bin*.

Spouštění programu

Pokud byl soubor zkopírovaný do adresáře s rodokmeny, spuštění programu provedeme příkazem:

./pedOrder

Zdrojové kódy souboru jsou uvedeny v adresáři *./PedOrder/libs* .

Datové soubory

Program pracuje s jedním vstupním souborem – rodokmenem, s libovolným názvem. Ukázkový soubor je pojmenován „rod.txt“.

Rodokmenový soubor

V tomto souboru jsou uvedeny rodokmenové záznamy studované populace.

Tento soubor má následující strukturu

Sloupec 1 – identifikační číslo jedince,

Sloupec 2 – identifikační číslo otce jedince

Sloupec 3 – identifikační číslo matky jedince

Pro správný běh programu je nutné z důvodu adresace jedinců, aby jedinci v rodokmenu byly vzestupně seříděny a překódovány vzestupně od nejstaršího jedince, který nese kód 1, po jedince nejmladšího (n). Podobně je to i u ostatních programů, které pracují s rodokmenovými soubory (například programy BLUP family (Misztal et al., 2002), DMU (Madsen a Jansen, 2008), či Endog (Gutiérrez and Goyache, 2005)). Neznámí jedinci jsou označeni kódem 0. Rodokmenový soubor je požadován ve volném formátu (oddělovač mezera) se vstupními hodnotami integer.

Ukázkové soubory jsou uvedeny v adresáři *./PedOrder/example* :

Rodokmenový soubor: rod.txt

Popis programu *pedOrder*

Po spuštění programu jsou po uživateli vyžadovány následující vstupy:

- název rodokmenového souboru,
- počet jedinců v rodokmenu,
- počet iterací – maximální počet běhů, které může program uskutečnit.

Po zadání uvedených vstupních informací se program spustí a během svého běhu vygeneruje následující soubory:

Soubor *order_ped.txt* – obsahuje původní rodokmenový soubor doplněný o hodnoty příslušné generace daného jedince.

Sloupec 1 – identifikační číslo jedince,

Sloupec 2 – identifikační číslo otce jedince,

Sloupec 3 – identifikační číslo matky jedince,

Sloupec 4 – příslušné číslo dané generace jedince.

V případě, že po ukončení iteračních běhů (dosažena maximální hodnota iterací) nedojde k ustálení hodnot generací jedinců, vygeneruje program soubor *errors.txt*. Tento soubor obsahuje výpis jedinců, u kterých nedošlo k ustálení hodnot generací. Tato skutečnost může mít dva důvody:

- Malý počet iterací, který neumožnil ustálení generací u všech jedinců. Z tohoto důvodu je doporučeno použít vyšší počet iterací cca 50.
- Soubor obsahuje zacyklený rodokmen – jedinec je sám jedním ze svých předků a z toho důvodu dochází stále k navyšování hodnoty generace u tohoto jedince a současně se zvyšují hodnoty generace i příbuzným jedincům.

Pro vyřešení druhého problému je možné použít několik přístupů:

- Všem jedincům vypsáním v souboru *errors.txt* dosadit nějakou maximální hodnotu generace.
- Všem jedincům vypsáním v souboru *errors.txt* změnit rodiče na neznámé předky (hodnota 0) a s takto opraveným rodokmenem zpustit program *PedOrder* znovu.
- Vyhledat mezi jedinci, uvedenými v souboru *errors.txt*, jedince, který způsobuje zacyklení rodokmenu a u tohoto jedince změnit jednoho rodiče na neznámého předka (hodnota 0). S takto opraveným rodokmenem zpustit program *PedOrder* znovu. Tento způsob je autory programu doporučován.