Characterizing cancer initiation by Moran Process

Jiantao Shi

26 November 2017

Abstract

It has been shown that the age incidence of chronic myeloid leukemia can be explained by a one-mutation Moran process model (Michor, Iwasa, and Nowak 2006). Moran process assumes the total number of cells is fixed over time, thus when a cell divides, another cell will be chosen by random to die. This model is in general useful to describe cancer initiation since the total number of pre-cancer cells tends to be stable. Here we implemented Moran process in Rcpp and apllied it to clone expansion of acute myeloid leukemia.

Contents

1 Intruduction

Mutations in oncogenes and tumor suppressor genes can increase the net reproductive rate of cells (fitness) and therefore provide a selective advantage for mutated cells. Most human cancers arise in epithelial tissues, which are organized into small compartments of cells (Nowak, Michor, and Iwasa 2003). Consider a compartment with N cells within a tissue, wild-type cells have fitness of 1, and mutant cells have fitness of r. In each cell division, a gene has a probability of u to mutate. In typical Moran process, any single time step consists of two elementary events: (1) a random cell is chosen for division proportional to its fitness, and (2) a random cell is chosen for elimination. Thus, total number of cells keep constant. Here we implemented Moran process in Rcpp and applied it to clone expansion of acute myeloid leukemia.

2 Implementation and example

There are 6 parameters in this model: total number of cells N, cell division rate τ (per day), relative fitness of mutant cells r ($r > 1$), mutation rate u, detection probability α given N total mutant cells, Age A (years). The core function is implemented in $MoranPorcess(N, A, \tau, r, u, \alpha)$. It returns values to indicate whether cancer is detected in an individual before age A (1 means detected). Detailed explanation of these parameters can be found at (Michor, Iwasa, and Nowak 2006).

```
library("Rcpp")
library("RcppArmadillo")
sourceCpp("../src/MoranProcess.cpp")

ls()
## [1] "MoranPorcess"
```

As shown above, `Rcpp` and `RcppArmadillo` are loaded to compile Rcpp source code, which results in `MoranPorcess`. Let's use parameters defined in (Michor, Iwasa, and Nowak 2006) as an example: $N = 2000$, $A = 80$, $r = 1.01$, $u = 10^{-6}$, $\tau = 40$, $\alpha = 2 * 10^{-5}$.

```
MoranPorcess(N = 2000, Age = 80, tau = 40, r = 1.01, u = 10^-6, alpha = 2*10^-5)
## [[1]]
## [1] 2000    0
##
## [[2]]
## [1] 0
```

This function returns a list with 2 elements. The first element is a vector of two integers: the number of wild-type cells and the number of mutant cells. The second element specifies whether cancer is detected before age A (positive integer means detected).

The example below simulates a population with 1000 individual and ask for the fraction for population has been detected to have cancer before age of 80. It take about 5 minutes to finish in a Macbook Pro.

```
    Instance = 0
    for(i in 1:1000){

        cat("Processing", i, "\n")

        temp = MoranPorcess(N = 2000, Age = 80, tau = 40, r = 1.01, u = 10^-6, alpha = 2*10^-5)[[2]]
        if(temp > 0)
            Instance = Instance + 1
    }

    Instance/1000
```

3 Acute myeloid leukemia

This mode can also be used to model acute myeloid leukemia (AML). A few studies have demonstrated that age incidence of mutation in acute myeloid leukemia driver genes increase over time (Jaiswal et al. 2014). Interestingly, most patients carry only one mutation in AML driver genes. This model was also applied and suggested that age incidence of AML can be explained by a one-mutation model.

4 Session info

```
sessionInfo()
## R version 3.4.1 (2017-06-30)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS High Sierra 10.13.1
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
```

```
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] RcppArmadillo_0.7.960.1.2 Rcpp_0.12.13
## [3] BiocStyle_2.4.1
##
## loaded via a namespace (and not attached):
##  [1] compiler_3.4.1  backports_1.1.1 magrittr_1.5    rprojroot_1.2
##  [5] tools_3.4.1     htmltools_0.3.6 yaml_2.1.14     stringi_1.1.5
##  [9] rmarkdown_1.6   knitr_1.17      stringr_1.2.0   digest_0.6.12
## [13] evaluate_0.10.1
```

References

Jaiswal, Siddhartha, Pierre Fontanillas, Jason Flannick, Alisa Manning, Peter V. Grauman, Brenton G. Mar, R. Coleman Lindsley, et al. 2014. "Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes." *New England Journal of Medicine* 371 (26): 2488–98. doi:10.1056/NEJMoa1408617.

Michor, Franziska, Yoh Iwasa, and Martin A. Nowak. 2006. "The Age Incidence of Chronic Myeloid Leukemia Can Be Explained by a One-Mutation Model." *Proceedings of the National Academy of Sciences* 103 (40): 14931–4. doi:10.1073/pnas.0607006103.

Nowak, Martin A., Franziska Michor, and Yoh Iwasa. 2003. "The Linear Process of Somatic Evolution." *Proceedings of the National Academy of Sciences* 100 (25): 14966–9. doi:10.1073/pnas.2535419100.