**Supplementary Table 3.** Conditional probability P(*bij*|*eijgi*) of read *bij* given true genotype *gi*, and read error *eij*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **True Genotype *gi*** | **Base Calling Error Event *eij*** | **Pr(*bij* =A)** | **Pr(*bij* =B)** | **Pr(*bij* =E)**[**b**](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3487130/table/tbl1/#tblfnb) |
| *gi* = AA[a](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3487130/table/tbl1/#tblfna) | *eij* = 0 | 1 | 0 | 0 |
| *eij* = 1 | 0 | 1/3 | 2/3 |
| *gi* = AB[a](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3487130/table/tbl1/#tblfna) | *eij* = 0 | 1/2 | 1/2 | 0 |
| *eij* = 1 | 1/6 | 1/6 | 2/3 |
| *gi* = BB[a](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3487130/table/tbl1/#tblfna) | *eij* = 0 | 0 | 1 | 0 |
| *eij* = 1 | 1/3 | 0 | 2/3 |

aAA, AB, BB: A allele homozygote, heterozygote, and B allele homozygote

bE: alleles other than A or B; assumes four possible alleles (bases)

(from Jun, G. et al. Detecting and Estimating Contamination of Human DNA Samples in Sequencing and Array-Based Genotype Data. *The American Journal of Human Genetics*, Vol. 91 839-848 (2012).)