Frequency distribution of the hereditary Alzheimer's disease-related genes seems to fit Poisson distribution, why?

Shufei Ge[1], Mingliang Cai[1,2] and Gang Pei[3,4 ✉]

Dear Editor,

We find, much to our surprise, that the percentage of mutation or duplication related to the early-onset Alzheimer's disease (EOAD, the well-established hereditary disease) reported in the World Alzheimer's Disease Report 2021[1] seems to fit a certain pattern. According to the report, *PSEN1* mutation accounts for the majority of EOAD gene mutations (43%), frequency of APP mutation represents 16%, and those of *PSEN2* and *APOE4* mutations are 6% and 9.12%, respectively, with an average of 7.56%. The next level is the rare mutations in genes *TREM2*, *SORLI*, *ABCA7* accounting for 1.17%, 1.42%, 1.33%, respectively[1–3], with an average of 1.31%. Now we can see that the reported percentages of the hereditary EOAD gene mutations with four different levels follow a similar pattern to Poisson distribution when the mean equals 1 (Fig. 1). In addition, we performed a goodness-of-fit test for the Poisson distribution hypothesis. With the total sample size of 170[2], the *P*-value of the test is 0.2415, which corroborates that the distribution of the reported EOAD-related mutation frequency fits a Poisson distribution with the mean being equal to 1.

An early study reports that the mutation frequencies in Big Blue mice also fit a Poisson distribution[4]. The number of emerging mutations per cell is also roughly Poisson distributed[5]. Hence, we wonder whether Poisson distribution is intrinsic to the 'frequency' distribution of EOAD-related mutations, or to that of any hereditary disease-related mutations. Or might it be true that Poisson distribution models frequency distribution of the gene mutations for any certain genotypes?

Interestingly, it is known that there are dozens of EOAD-related genes. We would like to hypothesize that there are some gene mutations occurring with even lower frequencies that might be close to level 5, namely ~0.31%, as predicted by Poisson distribution.

More intriguingly, the apparent fitness may provoke us to inquire into the biological meaning of the level 0. We do not know the answer at present, but we can still postulate that the level 0 could be a collection of many minor hereditary gene mutations, or the epigenetic modifications with hereditary properties, or even other hereditary risk factors unbeknownst to us yet, or the totality of the above.

We may take liberty to further speculate that for any tetraploid organism (human being is diploid), the relevant frequency would follow Poisson distribution with the mean of 2 (hexaploidy = 3, octoploid = 4, etc.).

Therefore, we would like to bring our hypothesis to community's attentions and hopefully to find more supporting or refuting evidence.

Correspondence: Gang Pei (peigang@tongji.edu.cn)
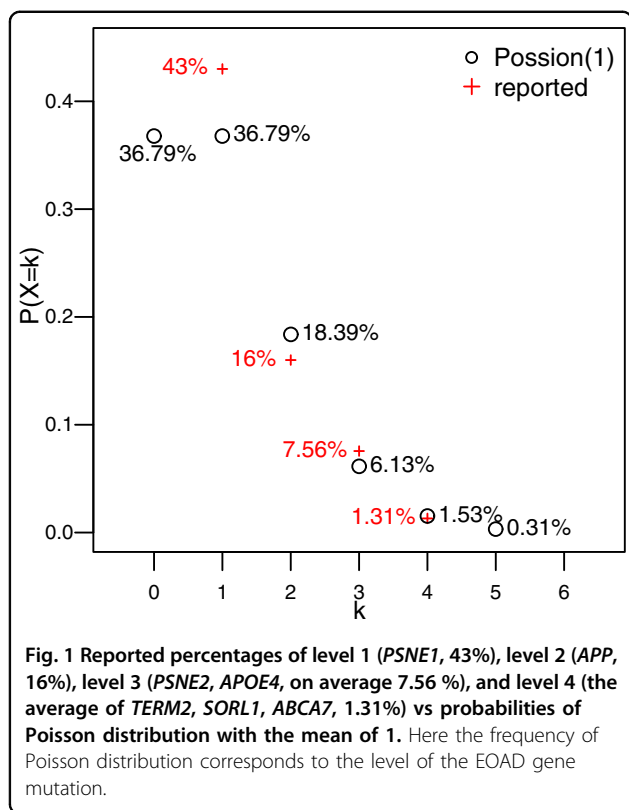[1]Institute of Mathematical Sciences, ShanghaiTech University, Shanghai, China
[2]Department of Mathematics, University of Miami, Coral Gables, FL, USA
Full list of author information is available at the end of the article

Fig. 1 Reported percentages of level 1 (*PSNE1*, 43%), level 2 (*APP*, 16%), level 3 (*PSNE2*, *APOE4*, on average 7.56 %), and level 4 (the average of *TERM2*, *SORL1*, *ABCA7*, 1.31%) vs probabilities of Poisson distribution with the mean of 1. Here the frequency of Poisson distribution corresponds to the level of the EOAD gene mutation.

Author details
[1]Institute of Mathematical Sciences, ShanghaiTech University, Shanghai, China. [2]Department of Mathematics, University of Miami, Coral Gables, FL, USA. [3]Shanghai Key Laboratory of Signaling and Disease Research, School of Life Sciences and Technology, Tongji University, Shanghai, China. [4]State Key Laboratory of Cell Biology, Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China

Author contributions
G.P. proposed the idea and advised on manuscript preparation. S.G. and M.C. wrote the manuscript. All authors reviewed and commented on the manuscript.

Conflict of interest
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Gauthier, S., Rosa-Neto, P., Morais, J. A. & Webster, C. World Alzheimer Report 2021: Journey through the diagnosis of dementia. Alzheimer's Disease International https://www.alzint.org/resource/world-alzheimer-report-2021 (2021).
2. Lanoiselée, H. M. et al. *APP*, *PSEN1*, and *PSEN2* mutations in early onset Alzheimer disease: A genetic screening study of familial and sporadic cases. *PLoS Med.* **14**, e1002270 (2017).
3. Bellenguez, C. et al. Contribution to Alzheimer's disease risk of rare variants in *TREM2*, *SORL1*, and *ABCA7* in 1779 cases and 1273 controls. *Neurobiol. Aging* **59**, 220.e1–220.e9 (2017).
4. Nishino, H., Schaid, D. J., Buettner, V. L., Haavik, J. & Sommer, S. S. Mutation frequencies but not mutant frequencies in Big Blue mice fit a Poisson distribution. *Environ. Mol. Mutagen.* **28**, 414–417 (1996).
5. Elez, M. et al. Seeing mutations in living cells. *Curr. Biol.* **20**, 1432–1437 (2010).