# Appendix

Figure S1 shows the probabilities of variants given components produced by the LDA model. It corresponds to Figure 3 presented in the article, which represents probabilities of components given variants. Figure S1 emphasizes the most frequent variants in the data, whereas Figure 3 emphasizes the most exclusive variants in the components. Please note that the probability scales of the figures do not match (0–1 in Figure 3, and 0–0.4 in Figure S1).

![Chart, bar chart

Description automatically generated]()

Figure S1. Highest conditional probabilities of variants given components.