Validation of *Kongoh* ver. 2.0.1

Kongoh ver. 1.0.1 was already validated by referring to the guidelines of the Scientific Working Group on DNA Analysis Methods (SWGDAM) [1, 2]. However, calculation results of *Kongoh* ver. 2.0.1 should be changed from those of ver. 1.0.1 because the drop-out model and some parameter values were modified. Therefore, the validation of *Kongoh* ver. 2.0.1 was performed on the basis of the SWGDAM guidelines. The validation metrics were as follows: sensitivity (i.e., LR of true contributors), specificity (i.e., LR of non-contributors), precision (i.e., reproducibility of LR values), case-type samples, and accuracy of calculations.

1. Testing set

Kongoh was validated using 27 two-person mixtures, 27 three-person mixtures, and 18 four-person mixtures. The information of DNA amount (ng) and mixture ratio are shown in Table 1. Six 2-person profiles included in PROVEDIt data sets were also used to validate case-type samples of DNA degradation and PCR inhibition. All DNA samples were amplified using the Identifiler Plus system with 28 amplification cycles. The PCR products were analyzed using an ABI 3130xl Genetic Analyzer with 10-s injection time, and the data are analyzed using GeneMapper^{TM} Software (Thermo Fisher Scientific). Analytical threshold was set at 30 relative fluorescence units (RFU). Thresholds for removing −1 backward stutter were not used. Further analyses were the same as those of *Kongoh* ver 1.0.1 [2].

Table 1. The information of experimental mixtures.

Mixture	No. of contributor sets	Mixture ratio	DNA amount (ng)
2-person	3	1 : 1	1, 0.5, 0.25
		3 : 1	
		9 : 1	
3-person	3	1 : 1 : 1	
		3 : 2 : 1	
		8 : 1 : 1	
4-person	2	1 : 1 : 1 : 1	
		4 : 3 : 2 : 1	
		7 : 1 : 1 : 1	

2. Sensitivity

We validated sensitivity using profiles of 27 two-person mixtures, 27 three-person mixtures, and 18 four-person mixtures, including samples of various DNA amounts and mixture ratios, without degradation and inhibition. We considered a true major contributor and a true minor contributor as POIs in each mixture, and calculated the LR values of each POI (Fig. 1). The LR values tended to decrease as the amount of DNA of the POI decreased. The LR values of the POI with more than 0.1 ng tended to be greater than 10,000, which can be considered as a "very strong evidence to support" the contribution by the POI [3]. In addition, there was no mixture that exhibited LR < 1 due to improving the allelic drop-out model. However, the LR values of POI with less than

0.1 ng DNA did not tend to be strongly supportive for the POI's contribution.

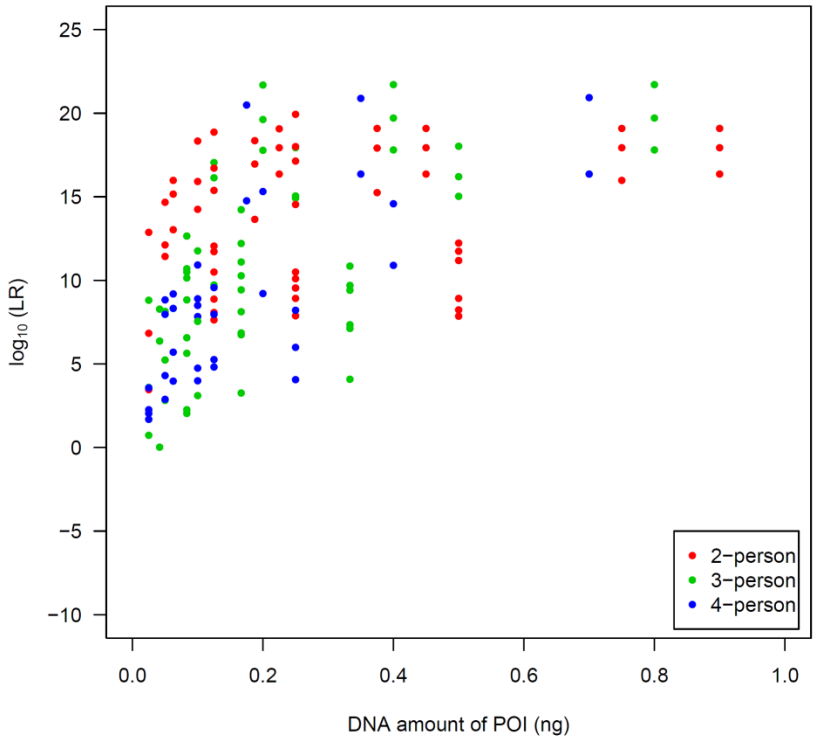

Fig. 1. Plots of POI's DNA amount vs LR values in true-contributor testing.

3. Specificity

For validating specificity, we computationally generated genotypes of 100 non-contributors based on Japanese allele frequencies [4], and then calculated the LR values assuming each non-contributor as a POI in 18 mixtures (Fig. 2). Most of the LR values were below one, that is, these LR values were not supportive for the POI's contribution to each mixture. In addition, all the mixtures except for a 4:3:2:1 mixture with 0.25 ng DNA indicated that the LR values of the non-contributor were smaller than those of true-contributor. However, the LR values of the non-contributor in *Kongoh* ver.

2.0.1 tended to be larger than those in *Kongoh* ver. 1.0.1 due to changing the allelic drop-out model. Especially in 0.25 ng mixtures with small amount of minor contributor's DNA, some LR values were greater than 1 (i.e., false inclusion). These LR values were near one and were not strongly supportive for the POI's contribution.

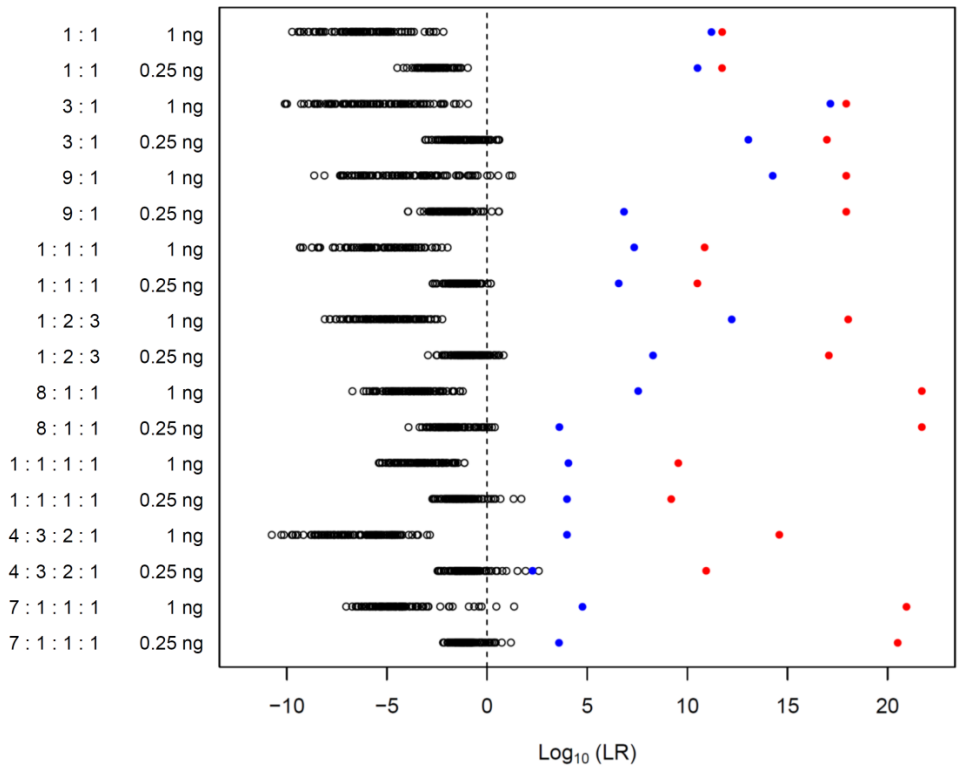

Fig. 2. LR values in non-contributor testing are shown in white circles. The red and blue plots represent LR values of the major and minor POIs, respectively.

4. Precision

For validating precision, we repeated the LR calculation of all 72 mixtures in true-contributor testing and 18 mixtures in non-contributor testing and compared the results of two runs (Fig. 3). The plots lie close to the $x = y$ diagonal line; therefore, there is little variation between each run. In true-contributor testing, the larger LR values were within twice the smaller LR values between two runs in 92.4% mixtures. The ratio of the larger LR to the smaller LR between two runs was at most 3.42 (i.e., LR = 3.52×10^{14} and 1.03×10^{14} in a 1:1 mixture with 0.5 ng DNA of the POI). In non-contributor testing, the difference of LR values between each run increased slightly. However, in 80.2% non-contributor testing, the larger LR values were within twice the smaller LR values.

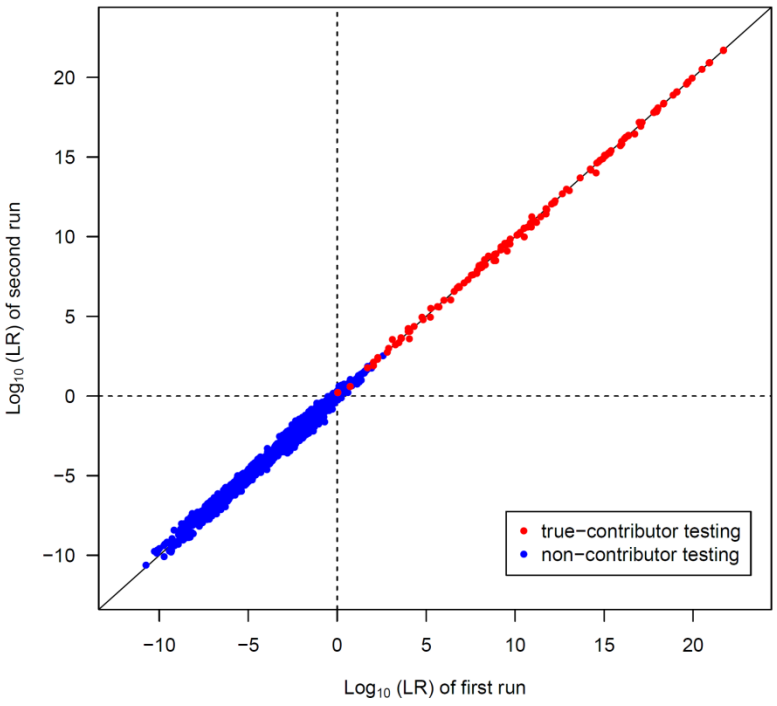

Fig. 3. LR values of two runs.

5. Case-type samples

The testing set that includes case-type samples of stutter, masked/shared alleles, and small amounts of DNA has high sensitivity, specificity, and precision as shown in Figs 1-3. We also interpreted the profiles of degradation and inhibition, which were obtained from the PROVEDIt dataset.

In samples with UV exposure, the LR values of the true major POI were strongly supportive of the contribution even in the samples with highly damaged DNA (Table 2). The LR values of the true minor POI were comparatively smaller than those of the major POI but greater than 10^5. DNA degradation parameters (d) reflect the degree of degradation and tend to be smaller with increasing DNA damage. The estimated d values demonstrated negative correlation with UV exposure times. Therefore, *Kongoh* was able to deal with peak heights appropriately in degraded samples.

In samples with inhibition through humic acid, the LR values of the true POI were strongly supportive of the contribution, even in the case of 2–7 drop-outs of the minor contributor. Unlike samples with UV exposure, peak heights did not correlate with molecular weights because inhibitors may act at each locus in the same manner regardless of the molecular weights. Therefore, the DNA degradation parameters were estimated as almost zero (i.e., no degradation).

Table 2. Calculation results of degradation and inhibition profiles.

DNA damage method	Estimated d value	LR (major POI)	LR (minor POI)
UV 15 min	−0.005	7.21×10^{19}	8.46×10^{7}
UV 60 min	−0.005	1.26×10^{11}	1.32×10^{5}
UV 105 min	−0.0075	1.52×10^{16}	1.27×10^{5}
Humic acid 15 µl	−0.0025	2.26×10^{22}	1.44×10^{13}
Humic acid 22 µl	0	1.83×10^{22}	5.49×10^{11}
Humic acid 35 µl	0	2.25×10^{22}	1.93×10^{12}

6. Accuracy of calculations

SWGDAM also recommends that studies should assess the accuracy of the calculations performed by the software. We confirmed the accuracy of each calculation performed by the *Kongoh* program, that is, the sampling from the probability distributions of biological parameters, gamma approximation of the expected peak heights, probability density of the observed peak heights given the expected peak heights, calculation of genotype frequencies, and calculation of LR. We also confirmed that the LR values for a major POI in mixtures with completely determined genotype of the POI through the peak height information (e.g., LR of a major POI in 9:1 mixture shown in Fig. 2) were the same as those of the reciprocal of the POI genotype frequency.

References

[1] Scientific Working Group on DNA Analysis Methods (SWGDAM). Guidelines for the Validation of Probabilistic Genotyping Systems. 2015.

[2] S. Manabe, C. Morimoto, Y. Hamano, S. Fujimoto, K. Tamaki, Development and validation of open-source software for DNA mixture interpretation based on a quantitative continuous model, PLoS One. 12 (2017) e0188183.

[3] I.W. Evett, G. Jackson, J.A. Lambert, S. McCrossan, The impact of the principles of evidence interpretation on the structure and content of statements, Sci Justice. 40 (2000) 233-239.

[4] K. Yoshida, K. Takahashi, K. Kasai, Allele frequencies of 15 loci using AmpFlSTR Identifiler Kit in Japanese population, J Forensic Sci. 50 (2005) 718-719.