Electronic Supplementary Material (ESI) for Analytical Methods

Supplementary information

1. The calibration sample sets and the prediction ones for the recognition models

For a certain genotype, all the measuring samples are obtained by the PCR amplification three times in the same amplification conditions, in which the two-thirds samples are collected as the calibration sample set and the rest samples as the prediction sample set. For other two genotypes, the measuring samples are also divided into calibration sample sets and prediction ones by the same method. And the detail result is shown in Table S1.

Table S1 The calibration sample sets and the prediction ones of genotypes 10-11,11-11 and 11-12

Genotype	Number	Sample set	Genotype	Number	Sample set	Genotype	Number	Sample set
10-11	1	prediction	11-11	1	prediction	11-12	1	prediction
10-11	2	calibration	11-11	2	calibration	11-12	2	calibration
10-11	3	calibration	11-11	3	calibration	11-12	3	calibration
10-11	4	prediction	11-11	4	prediction	11-12	4	prediction
10-11	5	calibration	11-11	5	calibration	11-12	5	calibration
10-11	6	calibration	11-11	6	calibration	11-12	6	calibration
10-11	7	prediction	11-11	7	prediction	11-12	7	prediction
10-11	8	calibration	11-11	8	calibration	11-12	8	calibration
10-11	9	calibration	11-11	9	calibration	11-12	9	calibration
10-11	10	prediction	11-11	10	prediction	11-12	10	prediction
10-11	11	calibration	11-11	11	calibration	11-12	11	calibration
10-11	12	calibration	11-11	12	calibration	11-12	12	calibration
10-11	13	prediction	11-11	13	prediction	11-12	13	prediction
10-11	14	calibration	11-11	14	calibration	11-12	14	calibration
10-11	15	calibration	11-11	15	calibration	11-12	15	calibration
10-11	16	prediction	11-11	16	prediction	11-12	16	prediction

10-11	17	calibration	11-11	17	calibration	11-12	17	calibration
10-11	18	calibration	11-11	18	calibration	11-12	18	calibration
10-11	19	prediction	11-11	19	prediction	11-12	19	prediction
10-11	20	calibration	11-11	20	calibration	11-12	20	calibration
10-11	21	calibration	11-11	21	calibration	11-12	21	calibration
10-11	22	prediction	11-11	22	prediction	11-12	22	prediction
10-11	23	calibration	11-11	23	calibration	11-12	23	calibration
10-11	24	calibration	11-11	24	calibration	11-12	24	calibration
10-11	25	prediction	11-11	25	prediction	11-12	25	prediction
10-11	26	calibration	11-11	26	calibration	11-12	26	calibration
10-11	27	calibration	11-11	27	calibration	11-12	27	calibration
10-11	28	prediction	11-11	28	prediction	11-12	28	prediction
10-11	29	calibration	11-11	29	calibration	11-12	29	calibration
10-11	30	calibration	11-11	30	calibration	11-12	30	calibration
10-11	31	prediction	11-11	31	prediction	11-12	31	prediction
10-11	32	calibration	11-11	32	calibration	11-12	32	calibration
10-11	33	calibration	11-11	33	calibration	11-12	33	calibration
10-11	34	prediction	11-11	34	prediction	11-12	34	prediction
10-11	35	calibration				11-12	35	calibration
						11-12	36	calibration
						11-12	37	prediction
						11-12	38	calibration

2. The discriminant models of the STR genotype based on PDV

The PDV method aims to find a balance between the separability and the stability with the model parameter λ between 0 and 1. For the discriminant analysis of three genotypes 10-11, 11-11 and 11-12, in order to achieve good between-class separability, at the same time take into account the prediction stability, the model parameter λ is optimized in the range (0, 1). The PDV models with λ-s of 10^{-4} and 10^{-6} were shown in Figs. S1 and S2, respectively.

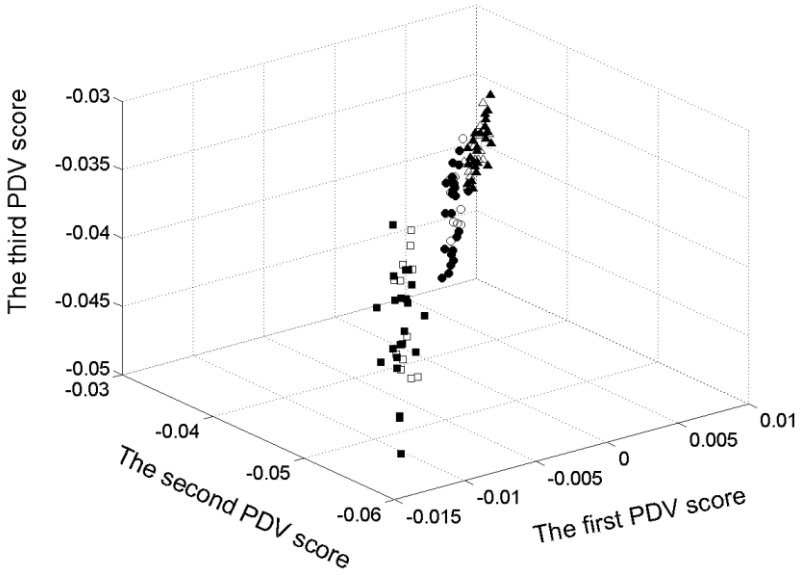

Fig. S1 The PDV model for genotypes 10-11, 11-11 and 11-12 with λ of 10^{-4}

("●" the calibration samples and "○" the prediction samples for genotype 10-11, "■" the calibration samples and "□" the prediction samples for the genotype 11-11, and "▼" the calibration samples and "∇" the prediction samples for the genotype 11-12)

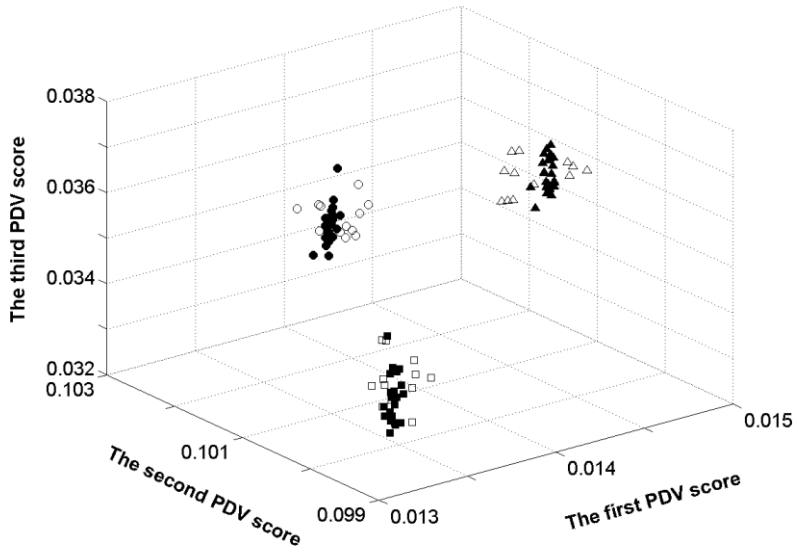

Fig. S2 The PDV model for genotypes 10-11, 11-11 and 11-12 with λ of 10^{-6}

("●" the calibration samples and "○" the prediction samples for genotype 10-11, "■" the calibration samples and "□" the prediction samples for the genotype 11-11, and "▼" the calibration samples and "∇" the prediction samples for the genotype 11-12)