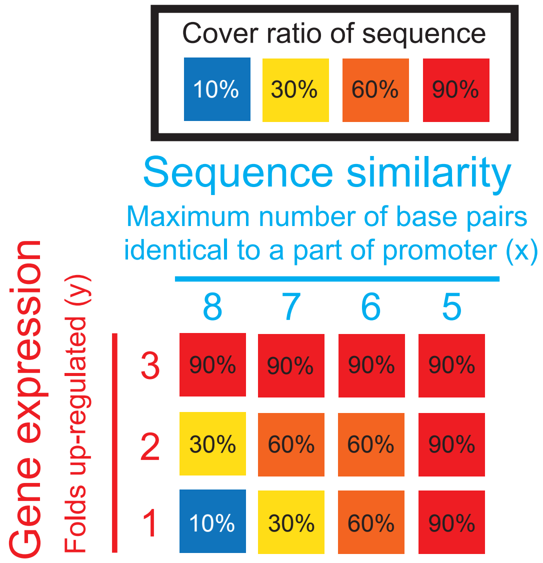
About MAMA score

The MAMA score was calculated from gene expression and sequence similarities. The following examples will help clarify how this score is related to gene expression and sequence similarities.

How to calculate MAMA score

CR = N(x|y)/N(y)

CR: Cover ratio
N(x|y): number of genes that have x bp of identical sequence to a candidate sequence and upregulated y-folds
N(y): number of genes upregulated y-folds



Cover ratio of sequence

🟦 10% 🟨 30% 🟧 60% 🟥 90%

Sequence similarity

Maximum number of base pairs identical to a part of promoter (x)

Gene expression

Folds up-regulated (y)

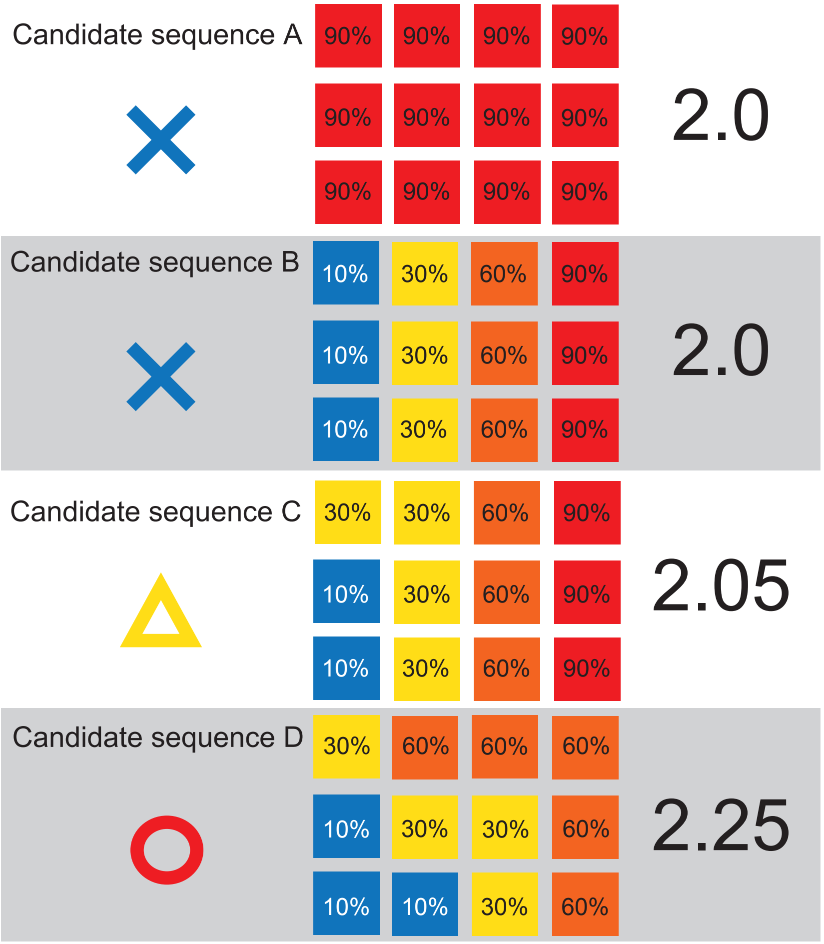
	8	7	6	5
3	90%	90%	90%	90%
2	30%	60%	60%	90%
1	10%	30%	60%	90%

$$^{*}\text{MAMA score} = \sum(CR \cdot x \cdot y) \,/\, \sum(CR \cdot x)$$

* In the paper, x was replaced with h_score, y was replaced with r_score, and each gene (1/1) was used instead of the cover ratio.

$$\{(0.9 \cdot 8 \cdot 3)+(0.9 \cdot 7 \cdot 3)+(0.9 \cdot 6 \cdot 3)+(0.9 \cdot 5 \cdot 3)$$
$$+(0.3 \cdot 8 \cdot 2)+(0.6 \cdot 7 \cdot 2)+(0.6 \cdot 6 \cdot 2)+(0.9 \cdot 5 \cdot 2)$$
$$+(0.1 \cdot 8 \cdot 1)+(0.3 \cdot 7 \cdot 1)+(0.6 \cdot 6 \cdot 1)+(0.9 \cdot 5 \cdot 1)\}$$
$$/\{(0.9 \cdot 8)+(0.9 \cdot 7)+(0.9 \cdot 6)+(0.9 \cdot 5)$$
$$+(0.3 \cdot 8)+(0.6 \cdot 7)+(0.6 \cdot 6)+(0.9 \cdot 5)$$
$$+(0.1 \cdot 8)+(0.3 \cdot 7)+(0.6 \cdot 6)+(0.9 \cdot 5)\} = 2.25$$

MAMA score and examples



					Score	Exsample
Candidate sequence A	90%	90%	90%	90%	2.0	Common sequence
×	90%	90%	90%	90%		
	90%	90%	90%	90%		
Candidate sequence B	10%	30%	60%	90%	2.0	Completely at random Sequences 8 bp identical are rare Sequences 5 bp identical are common
×	10%	30%	60%	90%		
	10%	30%	60%	90%		
Candidate sequence C	30%	30%	60%	90%	2.05	8 bp identical sequence is enriched in genes upregulated threefold when it is compared from that of onefold
△	10%	30%	60%	90%		
	10%	30%	60%	90%		
Candidate sequence D	30%	60%	60%	60%	2.25	8 bp identical sequence is enriched in genes upregulated threefold 7 bp identical sequence is enriched in genes upregulated two- and threefold 6 bp identical sequence is enriched in genes upregulated threefold
○	10%	30%	30%	60%		
	10%	10%	30%	60%		

MAMA scores become higher if identical sequences are enriched in upregulated genes.

Kakei et al. Additional file 2