Step-by-step calculation of the WTAA index in the *SIX3* Rescue assay

Classes

Three different phenotypes (Dead, Eyes and No eyes) are present. For our analyses, the phenotype Dead was excluded.

Proportion of Embryos with Eyes (PEE)

In general, PEE can be calculated as follows:

$$PEE_{i,m,j} = \frac{X_{E,m,j}}{X_{E,m,j} + X_{N,m,j}}$$

where $X_{E,m,j}$ and $X_{N,m,j}$ are, respectively, the number of embryos with Eyes (E) and No eyes (N) for the m allele (any allele) in the experiment j (any experiment). In our case, j varies between 2 and 7, and $m = 1,\dots,46$.

Two different PEEs were estimated: the global PEE, denoted as gP, and the experimental PEE, denoted as eP. While gP uses the total number of embryos through *all* the experiments for the estimation of the PEE, eP uses the information from single experiments only. For that reason, just one gP is calculated, whereas 2 to 7 ePs are calculated for each allele.

As an illustration, consider Table 1. Three different experiments were performed for this particular patient (allele *T316I*). For example, $eP_{T316I,1}$, the PEE in the first experiment is $eP_{T316I,1}$=14/(14+17)=14/31=0.4516. On other hand, the gP can be calculated as gP=(14+14+45)/(14+14+45+17+11+2)=73/103=0.7087. In addition, Table 2 shows the equivalency between alleles and patient number.

Wild Type Adjusted Activity (WTAA) index

The WTAA index is based-PEE measure of range (-∞, +∞), adjusted by *MO*, used to compare the activity of any allele with the activity of the *hSIX3* allele. Two different WTAA indices were estimated: the global WTAA index ($gWTAA$) and the experimental WTAA index ($eWTAA$). The $gWTAA$ index uses information about the total number of embryos in each class through *all* the experiments for a specific allele, while the $eWTAA$

considers the number of embryos in each experiment only. For the m allele, the $gWTAA$ index can be calculated as follows:

$$gWTAA_m = \frac{gP_m - gP_{MO}}{gP_{hSIX3} - gP_{MO}}$$

Note that by definition $gWTAA_{MO} = 0$ and $gWTAA_{hSIX3} = 1$. The $eWTAA$ in the experiment j for the m allele is given by the following expression:

$$eWTAA_{m,j} = \frac{eP_{m,j} - gP_{MO}}{gP_{hSIX3} - gP_{MO}}$$

Confidence Interval Calculation

To derive confidence intervals (CI) for the $gWTAA$ index, we performed bootstrap using the multinomial distribution as statistical model. By allele, a total of 10,000 simulations were performed over the total number of embryos in each class through *all* the experiments and the $gWTAA$ index was calculated for each of them. Then, by using the 2.5% and 97.5% percentiles from the $gWTAA$ empirical distribution, the 95% CI were estimated.

Plots

A new plot to represent the information is proposed (see Figure 1), where the X axis represents the patient number and the Y axis is the WTAA index, the blue horizontal line is the WTAA for the *hSIX3* allele, and the gray dotted vertical lines are divisors. Both the $gWTAA$ and $eWTAA$ indexes are represented by points (filled ones for $gWTAA$ and empty ones for $eWTAA$, respectively). Booststrap-based 95% CI intervals for $gWTAA$ are represented by solid vertical lines.

Software

Calculations, simulations and plot generation were done using R 2.7.0 Patched (R Development Core Team, 2008, http://www.R-project.org).

Tables

Table 1. Number of embryos in three experiments with the *T316I* allele (patient 46).

Experiment	**Phenotype**		**PEE**
	Eyes	**No Eyes**	
1	14	17	0.4516
2	14	11	0.5600
3	45	2	0.9574

Table 2. Coding of alleles by patient number.

Allele	**Patient**
G37C	1
G69D	2
A72P	3
V92G	4
A93D	5
D104G	6
I105V	7
W113X	8
W113C	9
S114L	10
E129X	11
S130X	12
R135P	13
V138D	14
404_407.dup	15
405_409.dup	16
406_407.dup	17
F157I	18
507.delG	19
A172V	20
H173P	21
Y174H	22
551del	23
556_557.dupl	24
582dupC	25
T202I	26
E207X	27
F213V	28
R218W	29
R218P	30

L226V	31
Q227P	32
P231R	33
232.234.del	34
A240K	35
Q241X	36
G244C	37
V250A	38
F254L	39
R257W	40
R257P	41
R258L	42
R262H	43
R269S	44
P297L	45
T316I	46

Figures

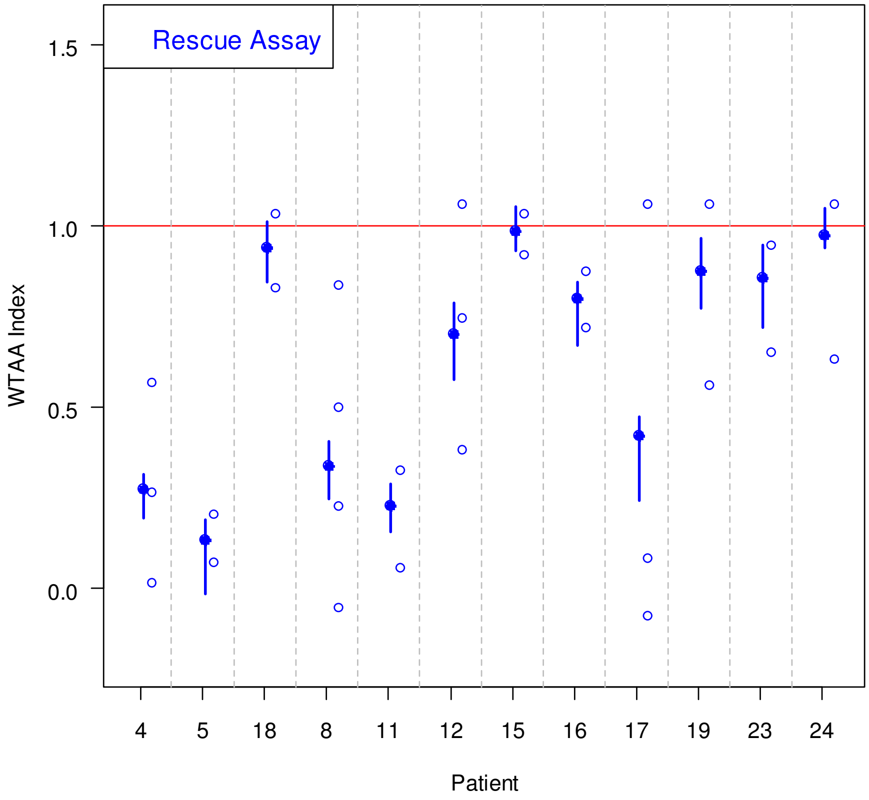


Figure 1. Proposed plot representing both the $gWTAA$ and $eWTAA$ index for 16 patients.

Step-by-step calculation of the Wild Type Adjusted Activity index in the *SIX3* overexpression assay

Classes

Four different classes were assigned to summarize the number of embryos with a particular phenotype: Class 1 (C1), Class 2 (C2), Class 3 (C3), and Wild Type (WT). See Fig 1 and Materials and Methods for details.

Phenotypic Index (PI)

PI is a continuous variable used to summarize the information by date and allele, in order to facilitate the comparison of the experiments. Conceptually, PI is similar to the expected value in statistics and can be calculated, generally speaking, using the expression:

$$PI_{i,m,j} = \frac{\sum_k w_i X_{i,m,j}}{\sum_k X_{i,m,j}}$$

where k is the number of classes, m the allele, w_i is the weight associated to the Class i, $i = 1,\ldots,k$, and $X_{i,m,j}$ is the number of embryos belong to the Class i for the allele m in the experiment j. For WT, C1, C2 and C3, w_i is defined as 0, 1, 2 and 3, respectively. In our case, $k = 4$, j varies between 2 and 10, and a total of $m = 1,\ldots,46$ alleles (patients) were studied.

Let us consider, for example, the number of embryos by phenotype in a specific experiment with the *hSIX3* allele. From Table 1, the number of embryos in WT, C1, C2, and C3 is 4, 3, 19 and 8, respectively. In addition, PI is ~1.911 and can be calculated as follows

$$PI = \frac{(0)(4) + (1)(3) + (2)(19) + (3)(8)}{4 + 3 + 19 + 8}$$

Wild Type Adjusted Activity (WTAA) index

The WTAA index is a PI-based measure of range [0, +∞), used to compare the activity of a particular allele with the activity of the *hSIX3* allele. Using the data set, two different WTAA indices were estimated: the global WTAA index ($gWTAA$) and the experimental WTAA index ($eWTAA$). While the $gWTAA$ index uses information about the total number of embryos in each class through *all* the experiments for a specific allele, the $eWTAA$ considers the number of embryos in each experiment only. For the m allele, the $gWTAA$ index can be calculated as

$$gWTAA_m = \frac{PI_m}{PI_{hSIX3}}$$

In addition, the $eWTAA$ in the experiment j (any experiment) for the m allele (any allele) is given by the following expression

$$eWTAA_{m,j} = \frac{PI_{m,j}}{PI_{hSIX3}}$$

To facilitate the understanding of both the $gWTAA$ and $eWTAA$ indices, we present some data in Table 2 and Table 3. Note that the number of embryos in Table 2 corresponds to the sum of embryos through *all* the experiments in Table 3. For example, the 75 embryos in C1 for the *hSIX3* allele (Table 2) represent 25 (C1 experiment 1) + 24 (C1 experiment 2) + 26 (C1 experiment 3). Also, from Table 2, $PI_{hSIX3} = 1.9394$ and $gWTAA_{G37C} = 1.3636/1.9394 = 0.7031$. Now, from the experiment 2 in Table 3, $PI_{hSIX3,2} = 1.9467$ and $eWTAA_{G37C,2} = 1.0232$. Table 4 shows the equivalency between alleles and patient number.

Confidence Interval Calculation

To derive confidence intervals (CI) for the $gWTAA$ index, we performed bootstrap using the multinomial distribution as statistical model. By allele, a total of 10,000 simulations were performed over the total number of embryos in each class through *all* the experiments and the $gWTAA$ index was calculated for each of them.

Then, by using the 2.5% and 97.5% percentiles from the $gWTAA$ empirical distribution, the 95% CI were estimated.

Plots

A new plot to represent the information is proposed (see Figure 1), where the X axis represents the patient number and the Y axis is the WTAA index, the blue horizontal line is the WTAA for the *hSIX3* allele, and the gray dotted vertical lines are divisors. Both the $gWTAA$ and $eWTAA$ indexes are represented by points (filled ones for $eWTAA$ and empty ones for $eWTAA$, respectively). Bootstrap-based 95% CI intervals for $gWTAA$ are represented by solid vertical lines. The colors red, orange and black represent WTAA<0.5, in 0.5-0.9, and >0.9, respectively.

Software

Calculations, simulations and plot generation were done using R 2.7.0 Patched (R Development Core Team, 2008, http://www.R-project.org).

Tables

Table 1. Number of embryos in an experiment with the *hSIX3* allele.

Class	**Embryos**
WT	4
Class 1	3
Class 2	19
Class 3	8

Table 2. Calculation of the gWTAA for the *hSIX3, G37C, G69D* and *A72P* alleles.

Allele	**Embryos by class**				**PI**	**gWTAA**
	C1	**C2**	**C3**	**WT**		
hSIX3	75	75	59	85	1.3673	1
G37C	55	60	105	131	1.3960	1.0210

G69D	76	80	52	111	1.2288	0.8987
A72P	57	43	108	148	1.3118	0.9594

Table 3. eWTAA calculation for the *hSIX3, G37C, G69D* and *A72P* alleles: (a) total number of embryos by experiment, (b) PI, and (c) eWTAA index.

(a)

Experiment	**Class**	**Allele**			
		hSIX3	***G37C***	***G69D***	***A72P***
1	C1	25	51	21	54
	C2	47	55	60	35
	C3	11	24	16	29
	WT	31	42	41	43
2	C1	24	3	31	3
	C2	4	2	8	3
	C3	38	78	33	77
	WT	9	38	11	42
3	C1	26	1	24	0
	C2	24	3	12	5
	C3	10	3	3	2
	WT	45	51	59	63

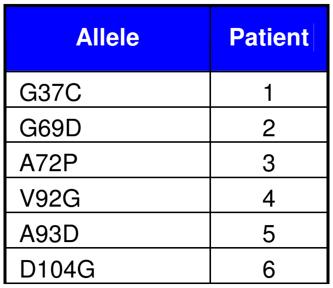
(b)

Experiment	**Phenotypic Index**			
	hSIX3	***G37C***	***G69D***	***A72P***
1	1.3333	1.3547	1.3696	1.3106
2	1.9467	1.9917	1.7590	1.9200
3	0.9905	0.2759	0.5816	0.2286

(c)

Experiment	**eWTAA**		
	G37C	***G69D***	***A72P***
1	0.9907	1.0016	0.9585
2	1.4566	1.2865	1.4042
3	0.2017	0.4254	0.1672

Table 4. Coding of alleles by patient number.



Allele	Patient
G37C	1
G69D	2
A72P	3
V92G	4
A93D	5
D104G	6

Figures

I105V	7
W113X	8
W113C	9
S114L	10
E129X	11
S130X	12
R135P	13
V138D	14
404_407.dup	15
405_409.dup	16
406_407.dup	17
F157I	18
507.delG	19
A172V	20
H173P	21
Y174H	22
551del	23
556_557.dupl	24
582dupC	25
T202I	26
E207X	27
F213V	28
R218W	29
R218P	30
L226V	31
Q227P	32
P231R	33
232.234.del	34
A240K	35
Q241X	36
G244C	37
V250A	38
F254L	39
R257W	40
R257P	41
R258L	42
R262H	43
R269S	44
P297L	45
T316I	46

Figures

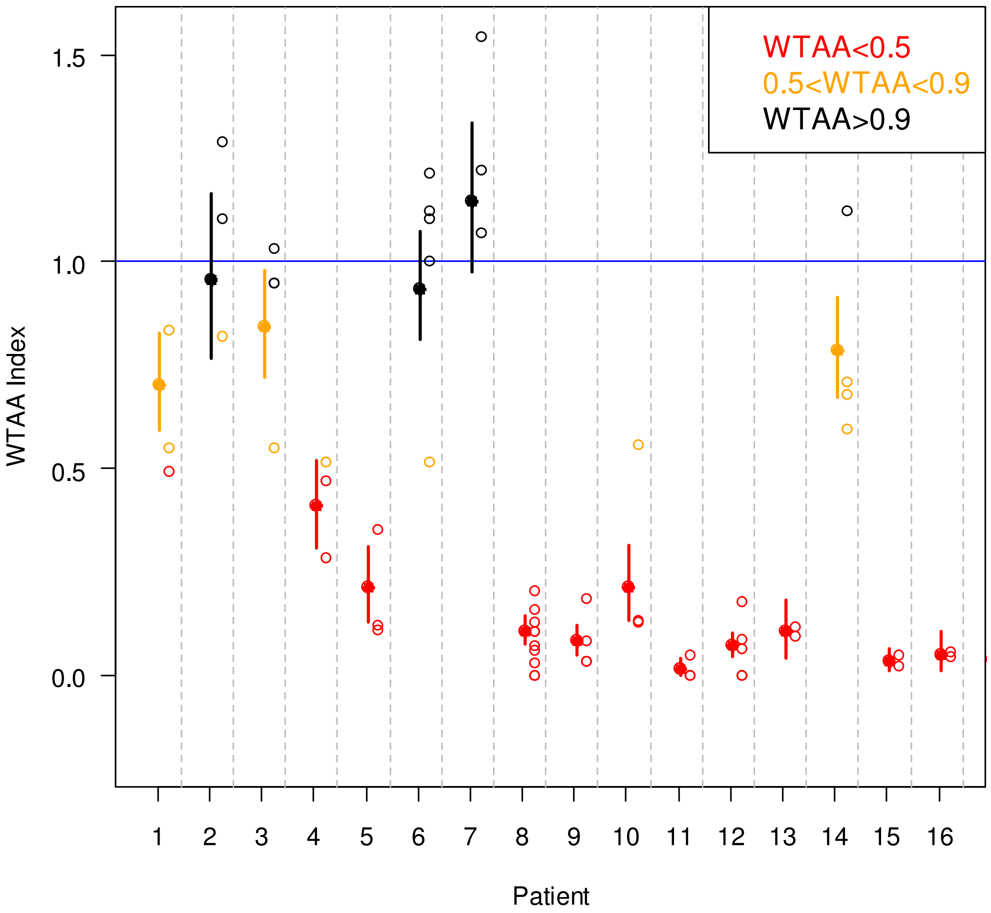


Figure 1. Proposed plot representing both the *gWTAA* and *eWTAA* index for 16 patients.