Classifying escape genes based on escape frequency threshold

A majority of the genes on the X chromosome are subject to X chromosome inactivation and only express one allele in each cell. Yet, ~10% of the genes consistently escape XCI, and another 12-20% of the genes are "variable escape" genes, which show considerable inter-individual differences in their escape status.

Most of the existing studies classify escape genes based on a threshold of the percentage of samples where each gene escapes XCI (which we call PE). For example, in the data set below *HCC1* escapes in 3/10 of the samples, resulting in PE of 30%.

gene	samp_state	sample	start	escape_freq
HCFC1	escape	HG00146	153947557	0.3
HCFC1	inactive	HG00163	153947557	0.3
HCFC1	escape	NA20502	153947557	0.3
HCFC1	inactive	HG00231	153947557	0.3
HCFC1	inactive	HG00235	153947557	0.3
HCFC1	escape	HG00258	153947557	0.3
HCFC1	inactive	NA20756	153947557	0.3
HCFC1	inactive	NA20769	153947557	0.3
HCFC1	inactive	NA20795	153947557	0.3
HCFC1	inactive	NA20819	153947557	0.3

The PE of a gene is used to classify its escape status. Genes which escape in nearly all samples are classified as "escape"; genes which escape in some but not all samples are "variable"; genes which escape in few or no samples are "inactive". **Escape Frequency Thresholds** are used to delineate each classification.

For example, the GEUVADIS lymphoblast study uses 25% and 75% thresholds.
- o %escape ≤ 25% → Silenced (S)
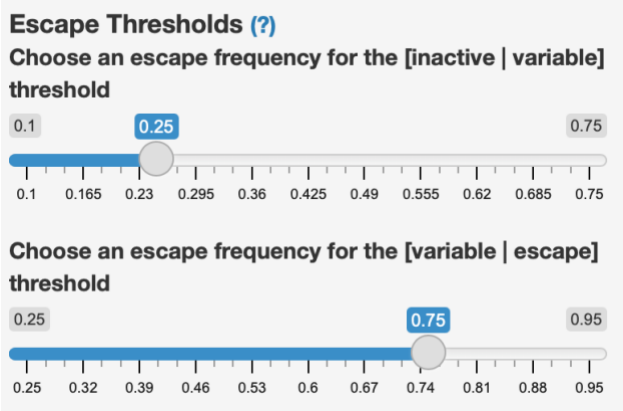- o 25% < %escape < 75% → Variable escape (VE)
- o %escape ≥ 75% → Escape (E)

Applying these thresholds to our example data set will classify *HCFC1* as "variable".

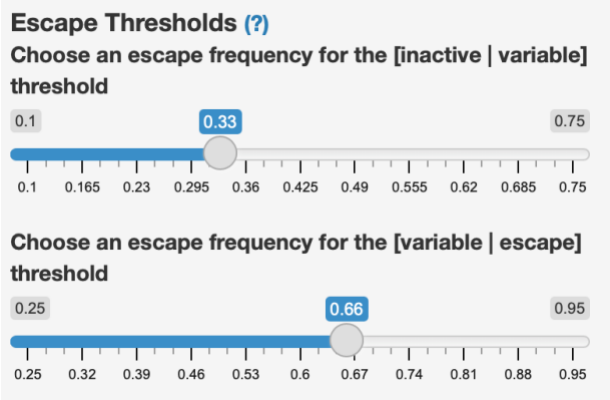| gene	samp_state	sample	start	escape_freq	state
HCFC1	escape	HG00146	153947557	0.3	variable
HCFC1	inactive	HG00163	153947557	0.3	variable
HCFC1	escape	NA20502	153947557	0.3	variable
HCFC1	inactive	HG00231	153947557	0.3	variable

CHANGING ESCAPE FREQUENCY THRESHOLDS

Different studies use different frequency threshold to classify escape, variable escape and inactivated genes. In order to harmonize results across studies and allow users to flexibly classify escape genes, X-Viz gives users the freedom to change the escape frequency thresholds.

The left side panel contains sliders which modify the lower and upper thresholds. These sliders are available only for studies which contain sample level information of escape status.

Escape Thresholds (?)

Choose an escape frequency for the [inactive | variable] threshold

0.1	0.25		0.75

0.1   0.165   0.23   0.295   0.36   0.425   0.49   0.555   0.62   0.685   0.75

Choose an escape frequency for the [variable | escape] threshold

0.25		0.75	0.95

0.25   0.32   0.39   0.46   0.53   0.6   0.67   0.74   0.81   0.88   0.95

Changing the escape thresholds will modify the classification sensitivity for the selected study. In our example data set, modifying the thresholds to 33% and 66% will re-classify the escape status of the *HCFC1* gene.
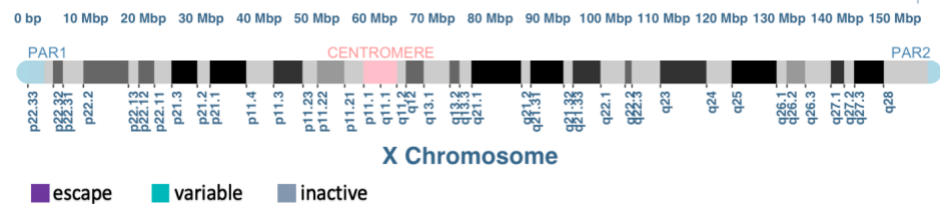
Escape Thresholds (?)

Choose an escape frequency for the [inactive | variable] threshold

0.1	0.33		0.75

0.1   0.165   0.23   0.295   0.36   0.425   0.49   0.555   0.62   0.685   0.75

Choose an escape frequency for the [variable | escape] threshold

0.25	0.66		0.95

0.25   0.32   0.39   0.46   0.53   0.6   0.67   0.74   0.81   0.88   0.95

gene	samp_state	sample	start	escape_freq	state
HCFC1	escape	HG00146	153947557	0.3	inactive
HCFC1	inactive	HG00163	153947557	0.3	inactive
HCFC1	escape	NA20502	153947557	0.3	inactive
HCFC1	inactive	HG00231	153947557	0.3	inactive

Modifying the escape thresholds will also update the gene position graphic. Note, this graphic is only available for studies which include positional data ("start" column).

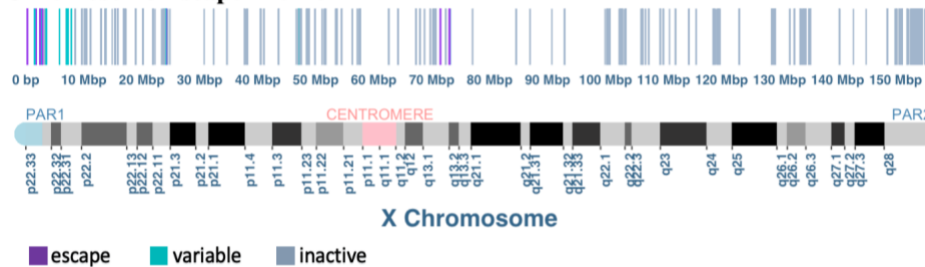**Graphic of 25% and 75% thresholds for *HCFC1*:**

X-Chromosome Escape Profile
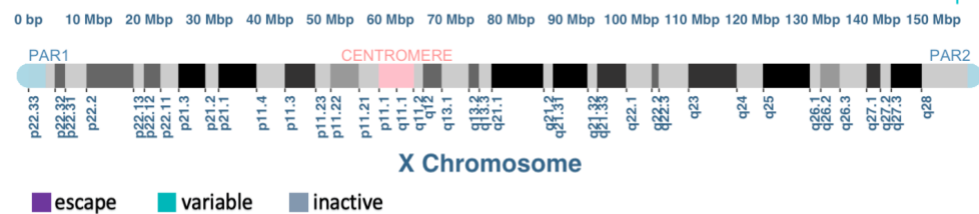
Graphic of 33% and 66% thresholds for *HCFC1*:

X-Chromosome Escape Profile

Graphic of 25% and 75% thresholds for *Tukiainen et al. GTEx study*:

X-Chromosome Escape Profile

Graphic of 33% and 66% thresholds for *Tukiainen et al. GTEx study*:

X-Chromosome Escape Profile