Computational analysis of disease-related mutation effects on transcription factor binding

TAMPERE UNIVERSITY OF TECHNOLOGY

Kirsti Laurila & Harri Lähdesmäki

Department of Signal Processing, Tampere University of Technology, Tampere, Finland

Motivation

Point mutation in gene regulatory regions

- Change in transcription factor (TF) binding
  - Binding affinity weakens (Figure 1a)
  - Binding affinity gets stronger (Figure 1b)

→ Altered gene expression levels

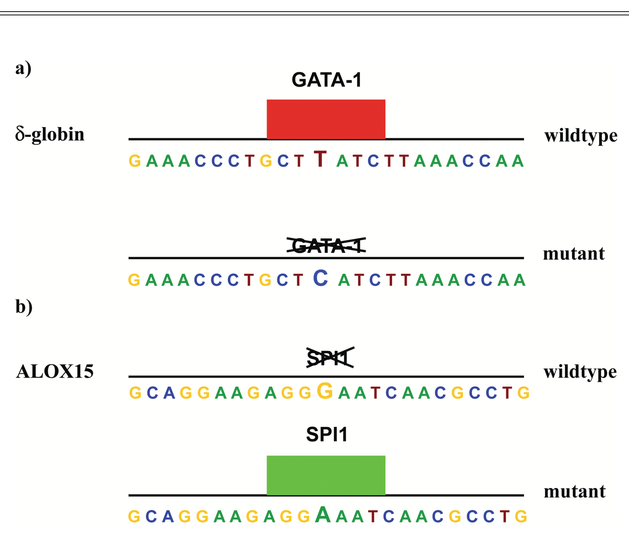

Figure 1: a) A point mutation in promoter of δ-globin gene weakens the binding affinity of TF GATA-1 [1]. b) A new TF binding site for TF SPI1 is caused by a point mutation in promoter of ALOX15 gene [2].

Methods

Experimentally verified mutations data set

- 21 mutations (Six of them presented in Table 1)
- Experimentally shown to affect TF binding

Mutations data set

- 474 mutations in 256 gene promoter regions
  - From Human Gene Mutation Database [3]

Binding affinity

- Likelihood ratio score based on position specific scoring matrixes (PSSMs) [4] and Markovian background model
- 343 different TFs, 496 PSSMs

Change in binding affinity

- Scores and p-values for wildtype and mutated sequences
- The p-value change is a score for the change in binding affinity
  - Negative → mutation weakens TF binding
  - Positive → mutation makes TF binding stronger

Mutation classes

- Dinucleotide steps RR, RY, YR, YY (R= A or G, Y= C or T)
- 8 mutation class where mutation is in the first nucleotide
- 8 mutation class where mutation is in the second nucleotide

Results

Experimentally verified mutations data set

- Part of results of experimentally verified mutations in Table 1

Mutations data set

- Estimated binding affinity limits for relevant mutations in the mutations dataset based on experimentally verified mutations
  - Change in affinity over 0.2 and the p-value for binding (wildtype or mutant) under 0.3
  - **OR** Change in affinity over 0.3
- Distribution of all changes in Figure 2a)
- Distribution of the relevant changes in Figure 2b)
  - Loss of binding affinity occurs more often than a new binding site is formed

Table 1: Experimentally verified mutations and their effect on TF binding. p-values are presented only for those PSSMs that show relevant changes. wt=wildtype, Δp-value=(p-value of wt) − (p-value of mutated sequence), mutation position is relative to TSS.

gene symbol	mutation	mutation position	TF	effect on binding	Δp-value	p-value of wt
ALOX	G→A	-292	SPI1	increase	0.356	0.592
HBD	T→C	-77	GATA1	decrease	-0.386	0.553
HBG2	C→G	-202	SP1	increase	0.274	0.540
HBG2	C→G	-202	SP1	increase	0.402	0.702
HBG2	C→G	-202	SP1	increase	0.658	0.861
HBG2	C→G	-202	SP1	increase	0.373	0.653
HBG2	C→G	-202	SP1	increase	0.206	0.420
PROC	T→C	-14	HNF-1	decrease	-0.216	0.265
UROS	C→A	-90	CP2	decrease	-0.207	0.143
UROS	C→A	-90	CP2	decrease	-0.274	0.164
UROS	T→C	-70	GATA1	decrease	-0.317	0.085
UROS	T→C	-70	GATA1	decrease	-0.206	0.038

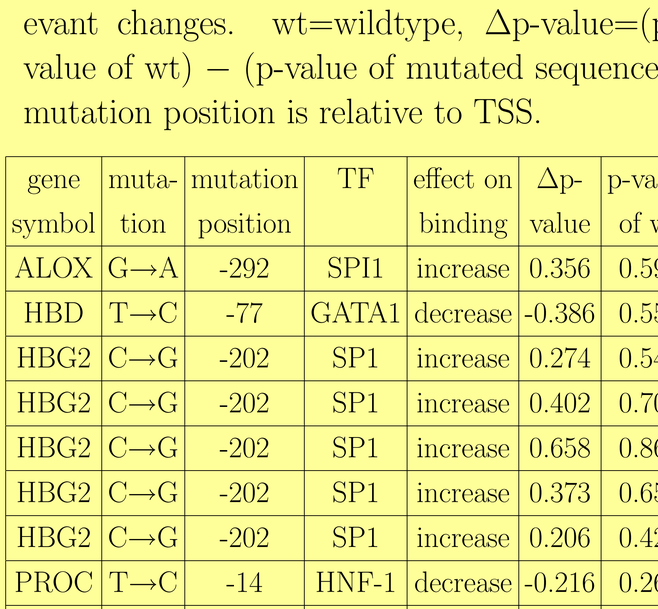

Figure 2: Distributions of the p-value changes. a) All changes. b) Only changes that exceeded the thresholds.

Mutation classes

- Different dinucleotide steps have different bendability [5]
  - YR steps are the most flexible steps
  - Steps AA, TT, GA, TC and AT are the most rigid ones
- DNA bending is often required in TF binding
- Can change in DNA bendability change TFs binding affinity? (See Table 2 and Figure 3)

Table 2: Effects of different mutation classes on TF binding. Underlined nucleotide means the mutated one.

mutation class	effect on binding
$\underline{R}R\rightarrow Y\underline{R}$, $\underline{Y}Y\rightarrow R\underline{Y}$, $R\underline{R}\rightarrow R\underline{Y}$, $Y\underline{Y}\rightarrow Y\underline{R}$	Binding affinity gets stronger more often than weakens (See Figure 3b)
$\underline{R}R\rightarrow \underline{R}R$, $\underline{Y}R\rightarrow \underline{R}R$, $Y\underline{R}\rightarrow Y\underline{R}$, $Y\underline{R}\rightarrow Y\underline{Y}$, $\underline{R}Y\rightarrow R\underline{Y}$, $\underline{R}Y\rightarrow \underline{Y}Y$ $R\underline{Y}\rightarrow R\underline{Y}$, $R\underline{Y}\rightarrow R\underline{R}$, $Y\underline{R}\rightarrow Y\underline{R}$	Binding affinity weakens more often than gets stronger (See Figure 3c)

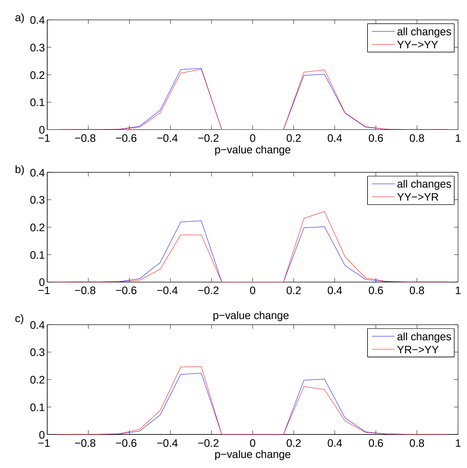

Figure 3: Distributions of the p-value changes in three different dinucleotide mutation types. a) YY→YY b) YY→YR c) YR→YY, Y=pyrimidine, R=purine.

Conclusions & Future Directions

- Regulatory mutation can change the TF binding significantly
- Number of false positives when using PSSMs?
- Using dinucleotide steps in modeling TF binding?
- Modeling binding of several TFs at the same time → protein-protein interactions

Acknowledgements

Financial support from Tampere Graduate School in Information Science and Engineering (TISE) is gratefully acknowledged. Work was also supported by the Academy of Finland, project nos 213462 (Finnish Programme for Centres of Excellence in Research 2006-2011), 106030, and 124615 and the Finnish Foundation for Technology Promotion.

References

[1] Matsuda M. et al. Blood, 80, 1347-51, 2006.

[2] Wittwer J et al. Hum Mutat. 27, 78-87, 2006.

[3] Stenson PD. et al. Hum Mutat, 21, 577-81, 2003

[4] Stormo GD. et al. Bioinformatics, 16, 1416-23, 2000

[5] Suzuki M. et al. Adv Biophys, 32, 53-72, 1996.