Bayesian Network of ENCODE TF-Target Interactions

Renqiang Min
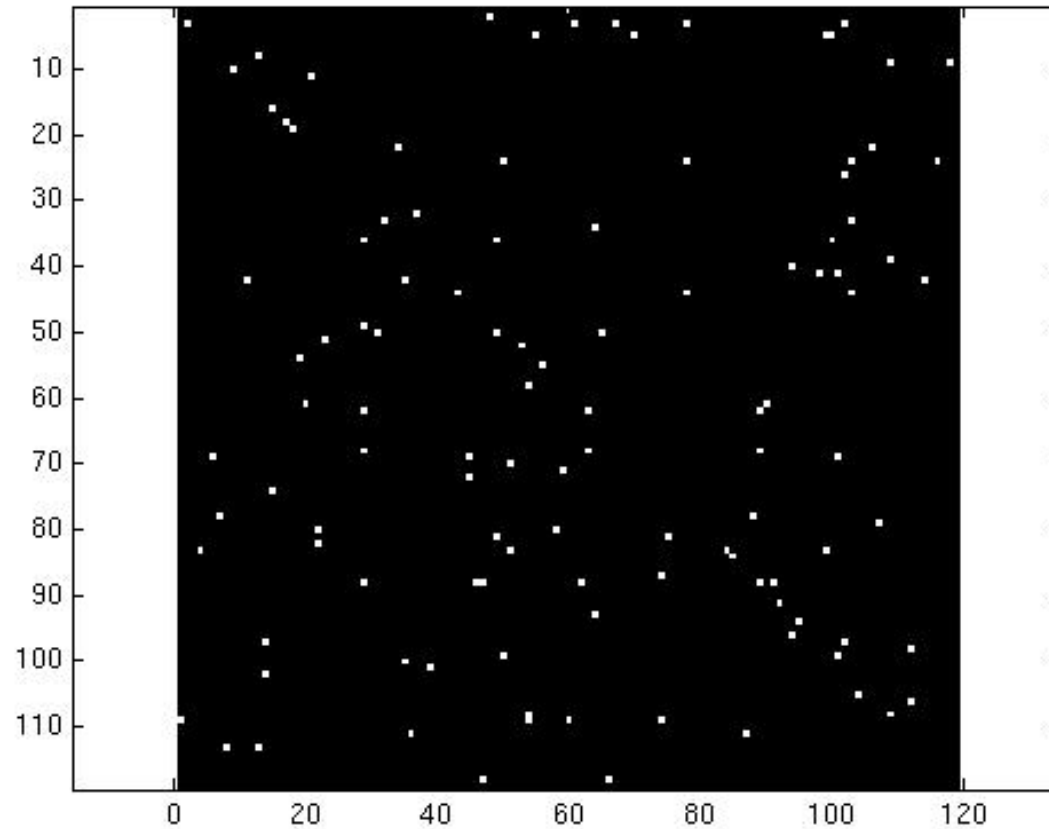
August 10, 2011

Bayesian Networks

- Using a graph to represent dependencies between variables, in which nodes correspond to variables and edges correspond to dependencies between variables

- Lack of edges represent conditional independency

- Main theorem: $p(V) = Product\_c\ p(V\_c|\ parents\ of\ V\_c)$; Conditioned on its parents, each variable is independent of all the variables except its children and descendants

Naïve Ways and Sophisticated Ways

- Given around 120 TFs and 9000 targets, we have an binary interaction matrix. View each target as a data point and each TF as a feature, and build a BN to model interactions between TFs to generate the whole data sets

- Consider protein-protein interactions between TFs, consider expression profiles of TFs and genes, and use these information to constrain structure learning

- Discriminative structure learning

Final Interactions between TFs

- 119 TFs

TFs with the largest TF Out-degrees

- SP1
- BCL11A
- BCLAF1
- EP300
- SMARCA4
- USF2
- GATA2
- HMGN3
- HNF4G
- JUND

- RXRA->SP1->[SP2 FOS IRF1 IRF3 NFYA SREBF1]
- BCL11A->[BATF NFKB1 PAX5 RXRA TCF12]
- BCLAF1->[MEF2A POU2F2 TAF7 TAL1]
- EP300 ->[JUND RXRA TCF4 ZNF143]
- SMARCA4->[BCL3 KAT2A SMARCB1 TAF7]
- [BRF1 GTF2F1 USF1]->USF2->->[ATF3 MAX NFE2 RAD21]

TFs with the largest TF In-degrees

- FOS
- JUN
- MAX
- RXRA
- SP2
- TBP
- TCF12
- TCF4
- USF2

- [GATA2 JUN NFYA PBX3 SP1]->FOS

- [GATA2 JUND SIRT6]->JUN

- [MYC USF1 USF2]->MAX

- [BCL11A EP300 HNF4G]->RXRA

- [NFYA PBX3 SP1]->SP2

- [HEY1 POLR3A TAF7]->TBP

TFs Without Downstream Targets

Bayesian Net

- BCL3
- BDP1
- CTBP2
- ETS1
- FOSL1
- STAT2
- TCF4
- TFAP2A
- TRIM28
- WRNIP1
- YY1

TFs in the top Layer in Bayesian Network and Hierarchies

- BCL11A
- BCLAF1
- BRF2
- CTCFL
- EP300
- ESRRA
- FOXA2
- HDAC2
- HEY1
- HMGN3
- HNF4G
- MAFF
- PBX3
- POLR3A

- POU5F1
- PPARGC1A
- SETDB1
- SIN3A
- SIRT6
- SIX5
- SMARCA4
- SRF
- STAT3
- SUZ12
- TFAP2C
- USF1
- XRCC4
- ZBTB33

- CCNT2
- CHD2
- CTBP2
- CTCF
- E2F1
- EBF1
- ELK4
- EP300
- ESR1
- GATA1
- GTF2B
- HDAC2
- HEY1
- HMGN3
- HNF4G
- IRF1
- JUN
- KAT2A
- MEF2C
- NANOG
- NFKB1
- NFYA
- NFYB

- NR2C2
- PBX3
- PPARGC1A
- RAD21
- RXRA
- SETDB1
- SMARCA4
- SMARCB1
- SPI1
- SREBF1
- SRF
- STAT3
- SUZ12
- TAF1
- TCF12
- TCF4
- TFAP2C
- TRIM28
- YY1
- ZBTB33
- ZBTB7A

Common Top-Layer TFs

- EP300
- HDAC2
- HEY1
- HMGN3
- HNF4G
- PBX3
- PPARGC1A

- SETDB1
- SMARCA4
- SRF
- STAT3
- SUZ12
- TFAP2C
- ZBTB33