

### Model Summary

- Total Training Genes – 287
- Number of Trees – 5000
- Number of Variables per Branch – 20
- Number of Pathways with correct genes - 6
- 1) KEGG Pathway 3010- Ribosome Synthesis
  - a) Model Statistics
    - i) Number of Training Genes: 35
    - ii) Number of Training Genes Correctly Identified: 21
      - (1) 10 genes classified in 4110
      - (2) All pathways had genes classified in 3010
        - (a) 4110 had 15 genes classified in 3010
        - (b) 190 had 10 genes classified in 3010
    - iii) Error: 42% (best of all pathways)
    - iv) Purity: 18% (117 genes placed in pathway)
    - v) Highest Importance Value: 3.44
  - b) Best Elements
    - i) FREAC\_3
      - (1) Genes 30
      - (2) Importance: 1.47
    - ii) Dof2
      - (1) Genes: 29
      - (2) Importance 1.32
    - iii) MNB1A
      - (1) Genes: 29
      - (2) Importance: 1.31
    - iv) Dof3
      - (1) Genes: 28
      - (2) Importance: 1.10
    - v) GATA3\_01
      - (1) Genes: 28
      - (2) Importance: 0.73
    - vi) PBF
      - (1) Genes: 27
      - (2) Importance: 0.98
    - vii) AP1\_Q2
      - (1) Genes: 14
      - (2) Importance: 1.71
    - viii) SP1\_01
      - (1) Genes: 12
      - (2) Importance: 1.35
    - ix) ELK1\_01
      - (1) Genes: 12
      - (2) Importance: 1.07
    - x) \*MZF1\_1\_4
      - (1) Genes: 11

- (2) Importance: 1.87
- xi) \*MZF1\_01
  - (1) Genes: 11
  - (2) Importance: 1.24
- c) Elements Ribosome had Highest Importance Value for
  - i) ELK1\_01
    - (1) Genes: 4
    - (2) Importance: 3.44
  - ii) TBP
    - (1) Genes: 4
    - (2) Importance: 2
  - iii) SAP\_1
    - (1) Genes: 7
    - (2) Importance: 1.93
  - iv) MZF\_5\_13
    - (1) Genes: 6
    - (2) Importance: 1.53
- d) About the Elements
  - i) AP1- Activator Protein 1
    - (1) Composed of Fos and Jun (Transfac)
      - (a) Fos/Jun heterodimer or Jun homodimer
    - (2) Jun
      - (a) KEGG Pathways (NCBI)
        - (i) 04662
        - (ii) 05210
        - (iii) 05120- Epithelial Cell Signaling in *Helicobacter pylori* infection
        - (iv) 04012- ErbB Signaling Pathway
        - (v) 04510- Focal Adhesion
        - (vi) 04912- GnRH Signaling
        - (vii) 04010
        - (viii) 05211- Renal Cell Carcinoma
        - (ix) 04660-
        - (x) 04620
        - (xi) 04310- Wnt Signaling
        - (xii) Reactome Event: Signaling in Immune System (168256)
      - (b) GO Annotation (NCBI)
        - (i) Cellular Process
        - (ii) Leading Edge Cell Differentiation
        - (iii) Regulation of Progression through Cell Cycle
    - (3) Fos is a regulator of cell proliferation, differentiation, and transformation (possibly apoptotic cell death) (NCBI)
      - (a) KEGG Pathways (NCBI)
        - (i) 04662- B Cell Receptor Signaling Pathway
        - (ii) 05210- Colorectal Cancer
        - (iii) 04010- MAP Kinase Signaling

---

\* Pathway had highest importance value for this element

- (iv) 04660- T Cell Receptor Signaling Pathway
- (v) 04620- Toll-like Receptor Signaling Pathway
- (b) GO annotation (NCBI)
  - (i) DNA methylation
  - (ii) Inflammatory response
  - (iii) Nervous System Development
- (4) Argument – Need many ribosomes for cells to undergo dramatic transformations via the production of many proteins
  - (a) May be a direct correlation between ribosome synthesis and the above processes, or ribosome assembly is another unique, unrelated job of this factor, if at all
- ii) Dof Transcription Factors
  - (1) Six Dof Genes: Dof1 (MNB1A), Dof2, Dof3, OBP1, NtBBF1, PBF
  - (2) Single Zinc Finger Transcription Factors that have homology to GATA transcription factors, but homology is considered insignificant
  - (3) Dof 1 and 2 may regulate tissue-specific and light-regulated transcription
    - (a) Light influences plant growth and development
    - (b) Light signals trigger leaf and chloroplast development and activation of photosynthetic genes
    - (c) Involved in expression of genes for carbon metabolism in maize
  - (4) PBF (Transfac0)
    - (a) Regulates cereal storage protein expression
    - (b) Stimulates heterologous storage protein in developing rice endosperm cells
  - (5) Dof transcription factors have only been found in plants
    - (a) Are the binding sites found evolutionary left-overs?
- iii) Elk Factors
  - (1) Transfac Database Information
    - (a) Mediates insulin-increased gene transcription
    - (b) Suppresses smooth muscle gene expression
  - (2) NCBI Database Information
    - (a) Nuclear target for the ras-raf-MAPK signaling cascade
    - (b) KEGG Pathways
      - (i) 05213- Endometrial Cancer
      - (ii) 04012
      - (iii) 04510
      - (iv) 04912
      - (v) 04910- Insulin Signaling
      - (vi) 04010
- iv) FREAC\_3
  - (1) Transfac Database Information
    - (a) Name: Fork Head Related Activator – 3
    - (b) Synonyms: FOXC1, FKHL7
    - (c) Gene: FOXC1
    - (d) Classification: fork head
  - (2) NCBI Database Information

- (a) Official Name: forkhead box C1
- (b) From the forkhead family of transcription factors
- (c) Function undetermined
- (d) Regulation of embryonic and ocular development
  - (i) Mutations cause glaucoma phenotypes
    1. primary congenital glaucoma
    2. autosomal dominant iridogoniodysgenesis anomaly
    3. Axenfeld-Rieger anomaly
  - (ii) Gene Ontology
    1. ureteric bud development
    2. paraxial mesoderm formation
    3. ovarian follicle development
    4. ossification
    5. odontogenesis (sensu Vertebrata)
    6. negative regulation of progression through mitotic cell cycle
    7. Notch signaling pathway
    8. artery morphogenesis
    9. blood vessel remodeling
    10. brain development
    11. camera-type eye development
    12. collagen fibril organization
    13. germ cell migration
    14. glycosaminoglycan metabolic process
    15. heart development
    16. in utero embryonic development
    17. kidney development
    18. lacrimal gland development
    19. mesenchymal cell differentiation
- v) GATA3\_01
  - (1) Transfac Database Information
    - (a) Name: GATA-binding Factor 3
  - (2) NCBI Database Information
    - (a) Synonyms: HDR (hypoparathyroidism, sensorineural deafness, renal dysplasia)
    - (b) Gene Ontology
      - (i) Anatomical structure morphogenesis
      - (ii) Cell fate determination
      - (iii) Defense response
      - (iv) norepinephrine biosynthetic process
      - (v) regulation of cytokine biosynthetic process
      - (vi) sensory perception of sound
      - (vii) sympathetic nervous system development
- vi) MZF – Myeloid Zinc Finger
  - (1) Transfac Database Information
    - (a) Negative regulator of CD34 and c-myb
    - (b) Possibly involved in hematopoietic development

- vii) SAP\_01
    - (1) Transfac Database Information
      - (a) SRF Accessory Protein
      - (b) Gene: ELK4
    - (2) NCBI Database Information
      - (a) Member of Ets family of transcription factors and TCF
  - viii) SP1\_01
    - (1) Transfac Database Information
      - (a) Name: Stimulating (specificity) Protein 1
      - (b) Phosphorylation activity
      - (c) Interacts with c-Jun, OCT-1, myogenin, YY1, NF-kappaB
      - (d) Cooperative binding with SREBP-1
      - (e) Multimerization leads to superactivation
        - (i) Effect depends on distance from TATA
      - (f) Synergistically cooperates with E2 and BPV promoters
      - (g) Important regulator of keratinocyte-specific gene expression in cooperation with AP-1 factors
        - (i) May give hint as to what factors regulate ribosome synthesis in different tissues
      - (h) Regulator of muscle-specific transcription in cooperation with myogenin (MyoD) and SRF
    - (2) NCBI Database Information
      - (a) KEGG Pathway
        - (i) 04350- TGF-beta signaling
  - ix) TBP
    - (1) Transfac Database Information
      - (a) Name: TATA binding protein
      - (b) Synonyms: TFIID, TFIIDtau
      - (c) Contacts: TFII125, PU.1, E1A, TAFII250
      - (d) Part of TFIID, SL1, and TFIIB
      - (e) Binds as monomer
      - (f) Bends DNA
        - (i) Enhanced by TFIIA
      - (g) TFIID250 mediates interaction with upstream factors
      - (h) Required for transcription by polymerase I, II, and III
- 2) KEGG Pathway 4070- PI Kinase
- a) Model Statistics
    - i) Number of Training Genes: 26
    - ii) Number of Training Genes Correctly Identified: 12
      - (1) 6 genes classified in 3010
      - (2) 5 genes classified in 4010
      - (3) Many pathways had genes classified in 4070
      - (4) 4110 had 8 genes classified on 4070
    - iii) Error: 54%
    - iv) Purity: 22% (55 genes placed in pathway)
    - v) Highest Importance Value: 4.46

- b) Best Elements
  - i) \*CAP\_01
    - (1) Genes: 26
    - (2) Importance: 4.46
  - ii) AP2alpha
    - (1) Genes: 22
    - (2) Importance: 1.72
  - iii) NF1\_Q6
    - (1) Genes: 11
    - (2) Importance: 2.61
  - iv) MNB1A
    - (1) Genes: 10
    - (2) Importance: 2.43
  - v) PBF
    - (1) Genes: 9
    - (2) Importance: 2.61
  - vi) Dof3
    - (1) Genes: 9
    - (2) Importance: 2.29
  - vii) MYB\_Q6
    - (1) Genes: 9
    - (2) Importance: 2.18
  - viii) Dof2
    - (1) Genes: 8
    - (2) Importance: 2.12
  - ix) \*AP1FJ\_Q2
    - (1) Genes: 6
    - (2) Importance: 3.77
- c) About the Elements
  - i) AP2alpha
    - (1) Transfac Database Information
      - (a) Synonyms: activator protein 2, ker1
      - (b) Gene: AP2-alpha, TFAP2A
      - (c) Dimerizes
      - (d) Interacts with c-myc, not max or mad
      - (e) Controls keratinocyte-specific gene expression
      - (f) Negatively regulates trans-activation by c-Myc
      - (g) Involved in gene response to PKA-PKC-mediated signals
      - (h) Involved in adrenomedullin expression
    - (2) NCBI Database Information
      - (a) GO Annotation
        - (i) Developmentally regulated
        - (ii) Ectoderm development
        - (iii) Signal Transduction
  - ii) CAP\_01

---

\* Pathway had highest importance value for this element

- (1) Transfac Database Information
  - (a) Cap Signal for Transcription initiation
- iii) MYB\_Q6
  - (1) Transfac Database Information
    - (a) Binding Factor: c-Myb
    - (b) Enhances DNA polymerase alpha expression in T lymphocytes
    - (c) Required for transition from Gap 1 to Synthesis stage in Cell Cycle
    - (d) Expression enhanced by c-Jun and JunD (but not by JunB) through an AP-1 element
  - (2) NCBI Database Information
    - (a) Official Full Name: v-myb myeloblastosis viral oncogene homolog (avian)
    - (b) Gene Ontology
      - (i) RNA splicing
      - (ii) Cell Cycle
      - (iii)mRNA processing
- iv) NF1\_Q6
  - (1) Transfac Information Database
    - (a) Synonyms: NF-1: Nuclear Factor 1, NFIC
    - (b) Gene: NFIC
  - (2) NCBI Database Information
    - (a) Gene: nuclear factor I/C (CCAAT-binding transcription factor)
    - (b) Synonyms: CTF, NFI, CTF5
    - (c) Gene Ontology
      - (i) DNA replication
      - (ii) Odontogenesis
- 3) KEGG Pathway 4110- Cell Cycle
  - a) Model Statistics
    - i) Training Genes: 34 genes
    - ii) Training Genes Correctly Classified: 9
      - (1) 65 genes classified to 4110
      - (2) All pathways have genes classified in 4110
      - (3) 15 genes were classified in 3010
      - (4) 8 genes were classified in 4070
      - (5) 10 Ribosome genes classified in 4110
      - (6) 7 Oxidative Phosphorylation genes classified to 4110
      - (7) 5 Polymerase genes
      - (8) 5 PI Kinase genes
    - iii) Error: 74%
    - iv) Purity: 14%
    - v) Highest Importance Value: 2.36
  - b) Best Elements
    - i) Elk\_1
      - (1) Genes: 14
      - (2) Importance: 1.78
    - ii) NFY\_Q6

- (1) Genes: 11
- (2) Importance: 1.21
- iii) Ahr\_ARNT
  - (1) Genes: 7
  - (2) Importance: 1.52
- iv) NF\_Y
  - (1) 1. Genes: 5
  - (2) Importance: 2.36
- v) \*MYCMAX\_01
  - (1) Genes: 4
  - (2) Importance: 1.67
- vi) \*PBX1\_01
  - (1) Genes: 3
  - (2) Importance: 1.20
- c) About the Elements
  - i) Ahr\_ARNT
    - (1) Transfac Database Information
      - (a) Binding Factors: Ahr/Arnt dimer
      - (b) Factor Name: Ahr (aryl hydrocarbon receptor)
        - (i) Synonyms: Ah receptor
        - (ii) Gene: AHR
        - (iii) Heterodimerization with ARNT required
        - (iv) Binds to xenobiotic response element
        - (v) NCBI Database Information
          - 1. ligand activated transcription factor
          - 2. regulates response to planar aromatic hydrocarbons
          - 3. regulates xenobiotic-metabolizing enzymes such as cytochrome p450
          - 4. ligands: aromatic hydrocarbons
          - 5. Gene Ontology
            - a. Apoptosis
            - b. Cell cycle
            - c. Response to stress
            - d. Response to xenobiotic stimulus
            - e. Signal transduction
    - (c) Factor Name: Arnt
      - (i) Transfac Database Information
        - 1. Synonyms: Ah receptor nuclear translocator, HIF-1beta, hypoxia-inducible factor 1beta
        - 2. heterodimer complexes with AhR or Sim
        - 3. required for AhR binding to XREs
        - 4. no ligand binding
        - 5. HIF1alpha/ARNT mediates gene responses to lowered oxygen levels
      - (ii) NCBI Database Information

---

\* Pathway had the highest importance value for this element



1. Ah receptor involved in induction of several enzymes for xenobiotic metabolism
  2. ligand-free AhR complexed to heat shock protein 90 (hsp90)
  3. binding of ligand, including dioxin, polycyclic aromatic hydrocarbons, results in transfer to nucleus
    - a. AhR and ligand bind to XREs
  4. ARNT forms complex with ligand and AhR
  5. ARNT required for AhR transcription
  6. ARNT identified as beta subunit of HIF1
  7. KEGG Pathways
    - a. 05211- Renal Cell Carcinoma
  8. Gene Ontology
    - a. Protein import into nucleus, translocation
    - b. Signal transduction
- ii) MYCMAX\_01
- (1) Binding Factors: c-myc/max heterodimers
  - (2) Transfac Database Information for c-Myc
    - (a) MAPK, GSK3, CKII, cdc2 phosphorylate c-Myc
    - (b) Interacts with p73alpha, AP2, pRb related protein p107 (inhibits c-myc)
    - (c) Causes release of cytochrome C from mitochondria and subsequent caspase-9 processing
    - (d) Proto-oncogene involved in cell proliferation control, may induce apoptosis
    - (e) May be involved in post-transcriptional regulation of rRNA metabolism
    - (f) Induces cyclin A and E expression and leads to a growth factor-independent association of cyclin A with E2F. Very early in cell cycle, myc acts to suppress expression of cyclin D1
    - (g) Can activate telomerase
      - (i) May contribute to its ability to promote tumor formation
    - (h) Negative auto-regulation
    - (i) Negative regulation by AP2, androgen receptor
    - (j) Overexpression of c-myc causes down-regulation of the thrombospondin-1 gene (an important negative modulator of tumor angiogenesis)
  - (3) NCBI Database Information for c-Myc
    - (a) Official Full Name: v-myc myelocytomatosis viral oncogene homolog (avian)
    - (b) Multifunctional, nuclear phosphoprotein plays a role in cell cycle progression, apoptosis, and cellular transformation
    - (c) Mutations, over-expression, rearrangement, and translocation associated with hematopoietic tumors, leukemias and lymphomas, including Burkitt lymphoma
    - (d) KEGG Pathway
      - (i) 05220- Chronic Myeloid Leukemia

- (ii) 05210- Colorectal Cancer
- (iii) 05213- Endometrial Cancer
- (iv) 04012- ErbB Signaling
- (v) 04630- Jak-STAT signaling
- (vi) 04010- MAP Kinase
- (vii) 05222- Small Cell Cancer
- (viii) 04350- TGF-beta signaling
- (ix) 05216- Thyroid Cancer
- (x) 04310- Wnt Signaling
- (e) Gene Ontology
  - (i) DNA fragmentation during apoptosis
  - (ii) Activation of pro-apoptotic gene products
  - (iii) Caspase activation
  - (iv) Cell cycle arrest
  - (v) Induction of apoptosis by intracellular signals
  - (vi) Iron ion homeostasis
  - (vii) Negative regulation of survival gene product activity
  - (viii) Positive regulation of cell proliferation
  - (ix) Regulation of apoptosis
  - (x) Release of cytochrome C from mitochondria
  - (xi) Response to radiation
- (4) Transfac Database Information for Max1
  - (a) Phosphorylation by CKII
  - (b) Max homodimers and Myc/Max and Myc/Mad heterodimers
  - (c) In high concentration, homodimers suppress c-myc by competition
  - (d) After myeloid differentiation, prefer to complex with Mad instead of c-Myc
- (5) NCBI Database Information for Max1
  - (a) Official Full Name: Myc associated factor X
  - (b) Dimers with Mad, Mxi1, Myc
  - (c) KEGG Pathways
    - (i) 04010- MAP Kinase
    - (ii) 05222- Small Cell Lung Cancer
- iii) NFY\_Q6
  - (1) Transfac Database Information
    - (a) Factor: NF\_Y, Nuclear Factor Y (Y-box binding factor)
    - (b) Synonyms: alpha\_CP1, CBF, CP1, NF\_Y; NF\_YA(h):NF\_YB(h):NF\_YC(h), trimer NF\_Y
    - (c) Three subunits
    - (d) Genes: CP1A, CP1B, CP1C
    - (e) Important regulator of MHC class II gene transcription
    - (f) Binds to X-Y-boxes cooperatively with RFX complex
    - (g) NCBI Database Information
      - 1. Genes NFYA, B, and C
      - 2. tissue-specific preferences for two isomeric forms
      - 3. KEGG Pathway

- a. 04612- Antigen Processing and Presentation
- iv) PBX1\_01
  - (1) Transfac Database Information
    - (a) Synonyms: pre-B-cell leukemia transcription factor 1a, Prl
  - (2) NCBI Database Information
    - (a) Gene Ontology
      - (i) C21-steroid hormone biosynthetic process
      - (ii) Cell differentiation
      - (iii) Embryonic development
      - (iv) Hindbrain development
      - (v) Sex differentiation
- 4) KEGG Pathway 190- Oxidative Phosphorylation
  - a) Model Statistics
    - i) Training Genes: 22 genes
    - ii) Training Genes Correctly Classified: 1
      - (1) 17 genes classified to 190
      - (2) Many pathways have 1 or 2 genes classified in 190
      - (3) 10, 20, 480, and 860 have 2 genes classified to 190
      - (4) 10 genes were placed in 3010
      - (5) 3 genes were classified in 4070
      - (6) 7 genes were classified in 4110
    - iii) Error: 95%
    - iv) Purity: 6%
    - v) Highest Importance Value: 3.06
  - b) Best Elements
    - i) AP2alpha
      - (1) Genes: 19
      - (2) Importance: 1.02
    - ii) GATA\_2
      - (1) Genes: 18
      - (2) Importance: 0.95
    - iii) DELTAEF1\_01
      - (1) Genes: 10
      - (2) Importance: 3.06
    - iv) deltaEF1
      - (1) Genes: 10
      - (2) Importance: 2.87
  - c) About the Elements
    - i) DelataEF1
      - (1) Transfac Database Information
        - (a) Species: chick, *Gallus gallus*
        - (b) Synonyms: delta-crystallin/E2-box factor 1, ZEB (human)
      - (2) Human Factor: AREB6
        - (a) Synonyms: Atpla1 Regulatory Element Binding Protein 6, deltaEF1, TCF8
        - (b) Gene: TCF8

- (c) Seven zinc fingers
  - (i) Extended form of Nil-2-a
  - (ii) Multiple conformations with respect to DNA binding
- (d) C-terminus zinc finger cluster: positive regulator
- (e) N-terminus zinc finger cluster: negative regulator
- (f) Interacts with general cofactor NC2
- (g) Represses interleukin-2 expression
- (h) Repressor of immunoglobulin heavy-chain enhancer
- (i) Represses promoter P1 of p73 gene in proliferating cells
- (3) NCBI Database Information
  - (a) Gene: ZEB1, zinc finger E-box binding homeobox 1
  - (b) TCF8 encodes human zinc finger transcription factor that represses T-lymphocyte-specific IL2 gene
  - (c) Gene Ontology
    - (i) Cell proliferation
    - (ii) Central Nervous System Development
    - (iii) Embryonic Morphogenesis
    - (iv) Immune Response
- ii) GATA 2
  - (1) Transfac Database Information
    - (a) Name: GATA-binding factor 2
  - (2) NCBI Database Information
    - (a) GATA transcription factors are candidate regulators of gene expression in hematopoietic cells
    - (b) GATA1: essential for normal primitive and definitive erythropoiesis
      - (i) Expressed at high levels in erythroid, mast cells, and megakaryocytes
    - (c) GATA2: expressed in hematopoietic progenitors including early erythroid cells, mast cells, and megakaryocytes
      - (i) Also in nonhematopoietic embryonic stem cells
      - (ii) In chicken erythroid progenitors, forced expression of GATA2 promotes proliferation at the expense of differentiation
    - (d) GATA3: expression restricted to T-lymphoid cells and some nonhematopoietic cell types, including embryonic stem cells
    - (e) Gene Ontology
      - (i) Cell fate determination
      - (ii) Cell maturation
      - (iii) Neuron Differentiation
      - (iv) Phagocytosis
      - (v) Pituitary gland development
      - (vi) Positive regulation of phagocytosis
- 5) KEGG Pathway 230- Purine
  - a) Model Statistics
    - i) Training Genes: 17 genes
    - ii) Training Genes Correctly Classified: 3
      - (1) 21 genes classified to 230

- (2) Many pathways have 1 or 2 genes classified in 230
- (3) 561, 3050, and 4070 have 2 genes classified to 230
- (4) 5 genes classified in 3010
- (5) 4 genes classified in 4070
- (6) 4 genes classified in 4110
- iii) Error: 82%
- iv) Purity: 14%
- v) Highest Importance Value: 3.42
- b) Best Elements
  - i) MYOD\_Q6
    - (1) Genes: 7
    - (2) Importance: 1.25
  - ii) \*NKX25\_01
    - (1) Genes: 6
    - (2) Importance: 1.76
- c) Other Elements
  - i) Egr3\_01
    - (1) Genes: 2
    - (2) Importance: 0.84
- d) About the Elements
  - i) Egr3\_01
    - (1) Transfac Database Information
      - (a) Name: early factor growth response gene 3 product
      - (b) Same binding as Egr-1 and -2
    - (2) NCBI Database Information
      - (a) An immediate-early growth response gene
      - (b) Induced by mitogenic stimulation
      - (c) Regulates genes that control biological rhythm
      - (d) Possible role in muscle development
      - (e) Gene Ontology
        - (i) Circadian rhythm
        - (ii) Muscle development
        - (iii) Neuromuscular synaptic transmission
        - (iv) Peripheral nervous system development
  - ii) MYOD\_Q6
    - (1) Transfac Database Information
      - (a) Factor Name: MyoD
      - (b) Description: myoblast determining factor
      - (c) Synonyms: MEF1, Myf-3
      - (d) Gene: MYOD1
      - (e) MyoD does not heterodimerize with other myogenic factors
      - (f) Positive cell specificity: myogenic cells
      - (g) Functional features: myogenic transcription factor
    - (2) NCBI Database Information
      - (a) Official Full Name: myogenic differentiation 1

---

\* Pathway received highest importance value for this element

- (b) Part of myogenic factors subfamily of basic helix-loop-helix family of transcription factors
  - (c) Regulates muscle cell differentiation by inducing cell cycle arrest, a prerequisite for myogenic initiation
  - (d) Involved on muscle regeneration
  - (e) Activates its own transcription
    - (i) May stabilize commitment to myogenesis
  - (f) Gene Ontology
    - (i) Cell differentiation
    - (ii) Multicellular Organismal Development
    - (iii) Myoblast Cell fate determination
    - (iv) Myoblast differentiation
    - (v) Protein amino acid phosphorylation
    - (vi) Striated muscle development
- iii) Nkx25\_01
- (1) Transfac Database Information
    - (a) Factor: Nkx2-5
    - (b) Role in maturation and morphogenesis of atrial and conduction tissue
    - (c) Atrial, ventricular and conotruncal septation, atrioventricular valve formation
    - (d) Mutations cause nonsyndromic, congenital heart diseases like cardiac malformations, atrial septal defects (ASD), atrioventricular conduction block ventricular septal defects (VSD), and tricuspid valve abnormalities (including Epstein's anomaly)
  - (2) NCBI Database Information
    - (a) Official Full Name: NK2 transcription Factor related, locus 5 (Drosophila)
    - (b) Homeobox-containing genes play critical roles in regulating tissue-specific gene expression essential for tissue differentiation, as well as determining the temporal and spatial patterns of development
    - (c) Drosophila homeobox-containing gene, tinman, expressed in developing dorsal vessel and in the equivalent of the vertebrate heart
      - (i) Mutations in tinman result in loss of heart formation in the embryo, suggesting that tinman is essential for Drosophila heart formation
    - (d) Abundant formation of Csx, the presumptive mouse homolog of tinman, is observed only in the heart from the time of cardiac differentiation
    - (e) CSX, the human homolog of murine Csx, has a homeodomain sequence identical to that of Csx and is expressed only in the heart, again suggesting important for human heart formation
    - (f) Gene Ontology
      - (i) Adult heart development
      - (ii) Cardiac muscle development
      - (iii) Embryonic heart tube development
      - (iv) Heart looping

- (v) Multicellular organismal development
- 6) KEGG Pathway 4010- MAP Kinase
- a) Model Statistics
- i) Training Genes: 9 genes
  - ii) Training Genes Correctly Classified: 1
    - (1) 2 genes classified to 4010
    - (2) 2 pathways have genes classified in 4010
    - (3) 1 gene classified in 561
    - (4) 4 genes classified in 3010
    - (5) 1 genes classified in 4070
    - (6) 3 genes classified in 4110
  - iii) Error: 89%
  - iv) Purity: 50%
  - v) Highest Importance Value: 9.80
- b) Best Elements
- i) \*AP2alpha
    - (1) Genes: 9
    - (2) Importance: 8.74
  - ii) \*c\_ETTS
    - (1) Genes: 7
    - (2) Importance: 2.58
  - iii) AP2\_Q6
    - (1) Genes: 6
    - (2) Importance: 9.80
- c) About the Elements
- i) AP2\_Q6
    - (1) Transfac Database Information
      - (a) Factor Name: AP-2, activator protein 2
      - (b) Website lists alpha and gamma forms
    - (2) NCBI Database Information
      - (a) Official Symbol: GTF3A
      - (b) Official Full Name: general transcription factor IIIA
      - (c) Synonyms: AP2, TFIIIA
  - ii) c-ETS (p54)
    - (1) NCBI Database Information
      - (a) Official Full Name: v-ets erythroblastosis virus E26 oncogene homolog 1(avian)
      - (b) KEGG Pathway
        - (i) 04320- Dorso-ventral axis formation
        - (ii) 05211- renal cell carcinoma
      - (c) Gene Ontology
        - (i) Immune response
        - (ii) Negative regulation of cell proliferation
        - (iii) Positive regulation of erythrocyte differentiation

---

\* Pathway received highest importance value for this element