Supplemental Fig S1: Web-logos of position weight matrices for 23 TFs investigated. Each weblogo is labeled with the TF name and TRANSFAC id.

Supplemental Fig S2: (A-B). Part of a sub-model taken from the *Interaction* model for CEBPB in H1hesc and Helas3 respectively. Each node in the tree is labeled with the TRANSFAC id, corresponding gene name and the threshold at which the feature is split. In each tree, one binding rule is highlighted: they are identical with respect to leading or increasing the probability of leading the binding probability of CEBPB.

(C). Accuracy (ROC-AUC) distribution for 6 choices of EMT feature sets. "<*" (">*") denotes significant difference (one sided Wilcoxon p-value < 0.05) between the two sets of performances and the direction (greater or less). The plot has same color coding as in Fig 1C-D, i.e. same color is used to denote the models using same features.

(D-F). ROC-AUC (D), number of sub-models (E), and fraction of sub-models (F) that included reference TFs, for different interaction depth of the models. Interaction depth defines maximum allowable features in a sub-model. The same colors denote model interaction depth for plots D-F.

(G-H). ROC-AUC (G), and number of sub-models (H) for different n.minobsinnode of the models. 'n.minobsinnode' denotes minimum number of observations made at each node while building the decision tree. Plots G and H use same the colors to indicate models using 'n.minobsinnode'.

(I-K). ROC-AUC (I), number of sub-models (J), and clustering consistency (K) of the sequences (percentage of sequence-pairs that fall in same clusters) for different values of the shrinkage parameter. Shrinkage indicates the learning rate of the model. Plots I-K the use same colors to denote models using same learning rate.

(L-M). Comparison of ubiquitous and cell type-specific sub-models: (L) Number of relevant features (i.e. features with non-zero importance in any cell type-specific model), (M) Standard deviation of feature importance for each sub-model, (N) Skewness of gene expression in each cells for the co-factors. Yellow and brown colors denote ubiquitous and cell type-specific sub-models respectively.

Supplemental Fig S3: Cluster membership matrix for k-nearest neighbor (KNN) clustering for k = 16. In each matrix, a row represents a cluster and a column represents a cell type. Elements in the matrix denote the number of sub-models in the cluster belonging to a specific cell type.

Supplemental Fig S4: Fraction of overlapped and non-overlapped sequences which fall in the same or different clusters. Dark orange represents the fraction of overlapped sequences falling in the same cluster, whereas light orange represents non-overlapped sequences. Dark purple represents the fraction of overlapped sequences falling in different clusters, with light purple non-overlapped sequences.

Supplemental Fig S5: Weblogos of the TRANSFAC ids with 85% similar to any zinger motifs.

Supplemental Fig S6: (A) Normalized within-cluster sum of squares stabilizes when the number of cluster is between 10 to 25; we choose k=16 (denoted by the vertical line in the closer view) as a representative for all TFs.

(B-C) Boxplot of *dh-ratio* (B), and *Hopkins statistic* (C) for 135 TF-cell pairs based on their sub-models, and pooled sub-models by TF. In (B), the horizontal line at Y=1 denotes the maximum limit of *dh-ratio*. In (C), the horizontal line at Y=0.6 denotes, current lowest value of *Hopkins statistic*.

(D) Same plot as in Fig 2C, except the sub-models are clustered by XY-Fused (XYF) self-organizing map. In plot (D-F), the 'blue' horizontal line denotes the coherence in 5% of the total multi-clusters.

(E-F) Functional and Expression coherence of sub-model clusters with expression threshold of log2CPM>=5, i.e. a gene is considered as on when the log2CPM>=5. ~40% (~18%) multi-cell clusters show higher expression-coherence (pathway-coherence). Dual coherence denotes both expression and pathway coherence. (E) is drawn for KNN (K-Nearest Neighbor) and (F) is drawn for XYF (XY fused network).

(G) Distance between binding sites and their nearest gene. (H) Distribution of enrichment scores of all relevant co-factor motifs (with nonzero feature importance), for each TF separately. The horizontal line at Y=1 denotes no enrichment/depletion, and the upper and lower dotted horizontal line denotes enrichment and depletion of 1.2 respectively.

Supplemental Fig S7: Cross-cell type performance matrix for *Interaction* and *Noninteraction* models. In each matrix, row represents the cell line used to build the model and column represents the cell line from which the test data is used. Diagonal elements are within cell type performance and only diagnal elements are colored according to the ROC-AUC to show the difference between *Interaction* and *Noninteraction* models.

Supplemental Fig S8: Relationship between model accuracy and sequence size. In each plot, color is used to indicate models from different cell lines.

Supplemental Fig S9: Same as Supplemental Fig S7, except the matrix is color coded according to the extent of symmetry of the non-diagonal elements. The symmetry is calculated by normalizing each row by the reference model (diagonal element).

Supplemental Fig S10: Motif usage for the reference TF in different cell types for the *NonInteraction* model. Y-axis denotes the feature importance of motif usage in the *NonInteraction* model. The sequence logos for the PWMs can be accessed from Supplemental Fig S1.

Supplemental Note 1. A model is built using Adaboost method where multiple classifiers are combined to represent the final output of the composite boosted classifier. In this approach, for each weighted bootstrap sample, a new sub-model (in our case a decision tree) is built and added to the model until no further improvement can be made.

In the *Interaction* model, the composite boosted model includes multiple decision trees, each of which captures a set of binding rules based on co-occurring motifs (potential interaction partners or co-factors) in the weighted training sequences bound by the reference TF. Each path from root to leaf in an estimated decision tree sub-model captures one such binding rule, asserting how a combination of motifs and along with their binding affinities relative to thresholds defined by the sub-model contribute to the target TF's binding. As an illustrative example, Fig 1B shows an arbitrarily selected sub-model of CEBPB in the Gm12878 cell line. Two of the binding rules are "presence of IRF8 with score greater than 2.08 *and* presence of NFATC4 with score of less than 2.3" - when these rules are met by the reference TF, CEBPB, is likely to bound. Whereas "presence of IRF8 with score greater than 2.08 *and* presence of NFATC4 with score of greater than 2.3" hinders CEBPB binding. While a general binding rule may be difficult to state concisely, it can be operationally defined in terms of a collective ensemble of cell type-specific binding rules. Each decision tree (a sub-model) operationally defines a binding rule, in terms of presence of specific motifs above/below a certain binding score. Furthermore, in general, the relative importance of features decrease with increasing depth of the node in the decision tree, with the first few levels contributing a substantial portion of the decision. Although a decision tree represents a statistical model for TF binding, by applying strict thresholds for motif scores and considering only the top few layers, in principal, a concise 'binding rule' can be derived, albeit, at a loss of information.

Each sub-model is built allowing interaction (tree) depth of 15. We found that 79% of all sub-models include the reference motif. However, this percentage increases up to 85% with increasing interaction depth (Supplemental Fig S2F) and no performance loss (Supplemental Fig S2D). The sub-models without a reference motif may be explained by the possible absence of the reference motif sequence from the training set due a sequence-length restriction, PWM match threshold, indirect binding, or by other unknown confounders. Notably, by virtue of physical space, the number of non-overlapping features fit in the sequence of restricted length should be limited, e.g. with average size (8bp) of PWM the 15 features need at least 120bp. However, we have ascertained that only 0.8% of all possible single paths (encompassing only ~13% of all the sub-models) have more than 12 features. Thus, in almost all cases, the features fit in a 100bp physical space (12*8=96).

We clustered the sub-models based on feature importance, meaning the contribution of each co-factor in the set of binding rules specified by each decision tree. Therefore, by design, sub-models, common across cell types, will have increased similarity in the set of co-factors and the weight of their contribution whereas cell type-specific sub-models will either have different sets of co-factors, or similar sets of co-factors but with different contribution. For example, none of the CEBPB sub-models have the rule: "presence of IRF8 *and* presence of NFATC4 leads to the binding of CEBPB" (highlighted in Fig 1B) except in Gm12878. On the other hand, the following rule exists in multiple sub-models of all cell types except Gm12878 generated sub-models: "presence of CEPBE and one of the reference PWM of CEBPB increases the binding probability of CEBPB". In Supplemental Fig S2A-B the corresponding rules are highlighted. Overall, when we looked at the ubiquitous sub-models and cell-specific sub-models, we found that ubiquitous sub-models contribute more co-factors than cell type-specific sub-models (Supplemental Fig S2L). This leads to cell-specific sub-models having a more skewed feature importance than ubiquitous sub-models (Supplemental Fig S2M).

Furthermore, the co-factors contributed by cell-specific sub-models exhibit a slightly more skewed gene expression across cell types than those contributed by ubiquitous sub-models (Supplemental Fig S2N, see Methods for details).

Supplemental Note 2. Each cluster of sub-models can itself serve as a composite, or ensemble, classifier. We determined a cluster-specific score for each TF-bound sequence based on these new cluster-based ensemble classifiers and assigned each sequence to one or more sub-model clusters based on this score (see Methods). Independently, for each TF, we partitioned all bound sequences into those that are bound uniquely in a cell type and those that are bound in multiple cell types. In general, if the clustering of sub-models in different cell types is simply due to sequence sharing then we expect to see a large fraction of overlapping sequence pairs, and not the cell type-specific sequence pairs, assigned to same cluster. We trained EMT using 75% of the sequences in each cell type dataset followed by clustering, and assessed the aforementioned fractions for the remaining 25% of the sequences to avoid training bias. As shown in Supplemental Fig S4, we expect pairs of overlapping sequences to be assigned to the same cluster and hence the size of dark orange box (same cluster) is greater than dark purple box (different cluster). However, many pairs of non-overlapping sequences are also assigned to the same cluster (light orange). We conducted a chi-squared test to assess whether the proportion of non-overlapped sequence pairs assigned to the same cluster is smaller than expected, indicating that co-clustering is driven by sequence overlap. For each cluster, we obtain the proportion of overlapping and non-overlapping sequence pairs and computed the expected proportion from the overall proportion of overlapping vs. non-overlapping sequence pairs. We conducted one chi-squared test per TF using data pooled from all clusters and all cell types, there was no evidence for depletion of non-overlapped sequence pairs assigned to the same cluster (all P-values > 0.05). These results suggest that co-clustering of sub-models across cell types are not simply due to sequence overlap, but rather, represent shared binding rules.

Supplemental Note 3. In clustering the sub-models, our goal was not to find the precise number of distinct binding rules, but rather to assess the modularity and sharing of binding rules across cell types. That's why we decided to choose k in such a way that the coherence among the sub-models in the same clusters is still detectable (i.e. k, not too high) while still revealing the cross-cell type sharing (i.e., k not too low). We checked the value of within-cluster sum of squares (normalized by the cluster-size) for different cluster sizes (Supplemental Fig S6A) (Wien 2002). For some TFs the suggested clusters seem ~15 (e.g. CTCF), for others ~20 (e.g. FOS), and in extreme cases the desired number of clusters seems to be more than 30 (e.g. ATF3, MYC). Based on these results a cluster size ranging from 15 to 25 seemed a reasonable choice. As a compromise across TFs and to make the analyses comparable we selected k=16 for all TFs.

Supplemental Note 4. We collected 22 position frequency matrices for the zinger motifs reported in (Worsley Hunt & Wasserman 2014). We identified the corresponding TRANSFAC id by matching the PWM-similarity by TFBSTools R package (Lenhard & Wasserman 2002). Allowing 90% (85%) PWM-similarity gave us 16 (42) TRANSFAC ids as zinger motifs; Supplemental Fig S5 lists all the zinger TRANSFAC ids. We found that only 5.5% (14%) of the identified the co-factors are zinger motifs suggesting that these motifs have little impact on the models. Moreover, we checked the clustering pattern of the sub-models after removing the

zinger motifs and found that the *sparsity* of the cluster-membership matrix is highly correlated with the original clustering pattern (spearman correlation = 0.96, p.value = 2.4 x 10^{-13}). This suggests that our overall findings are not affected by the zinger motifs.

Supplemental Note 5. It is possible that *EMT*s can falsely yield multiple sub-models, even in absence of heterogeneity, and those sub-models can be falsely clustered. We ascertained heterogeneity across sub-models for a TF from multiple cell types using a *Duda-Hart test* (Duda et al. 2001) and assessed the clustering tendency of the sub-models in the *d*-dimensional feature space using *Hopkins statistics* (Jain & Dubes 1988). The *Duda-Hart* test verifies whether or not a set of data points should be split into two clusters from the estimate of within-cluster sum of squares for all pairs of clusters versus overall sum of squares; the ratio of the two sum of squares is quantified as the *dh-ratio*; the smaller the value, the greater the clustering. On the other hand, the *Hopkins statistic (H)* compares the nearest neighbor distribution for a random set of points to the same distribution for the clustered sub-models (see Methods). A value close to 0.5 indicates the sub-models are random sets of points with no clustering, a value close to 1 indicates that they form cohesive clusters. Supplemental Fig S6B-C summarize the *dh-ratio* and *Hopkins statistic* respectively for 135 TF-cell pairs based on sub-models of TF-cell type pair, and for each TF after gathering all sub-models under a TF. We found that in all cases the *dh-ratio* is less than 1, and the *Hopkins statistic* > 0.5, consistent with heterogeneity; all tests rejected homogeneity (p.value <0.001). Together, the *Duda-Hart test* and *Hopkins statistic* strongly suggest that the sub-models are distinct and cluster-able, i.e., TF binding rules are heterogeneous and partly shared across cell types.

Supplemental Note 6. Supplemental Fig S6H shows the distribution of enrichment scores for the co-factors identified per TF. Except CTCF, the minimum of median enrichment score is ~1.2. If we choose a cutoff greater than 1.2, we might lose true positive co-factors for TFs like NRF1, REST, TBP etc. On the other hand, a lower threshold will likely yield many false positive co-factors for the other TFs.

Supplemental Note 7. The enriched GO terms (only biological processes at <=10% false discovery rate) for the cell type-specific co-factors are listed in Supplemental Table 9. Here we discuss the literature evidence supporting the TF functionality in different tissues related to the enriched terms for some of the TFs studied. For other TF's, based on our limited literature survey, we did not find a compelling support for tissue-specific functions of the TF. The following should be considered a selected sampling and an absence of support below should not necessarily be considered as a contradiction.

1. **BHLHE40** – BHLHE40 is known to be associated with many biological processes including circadian rhythm (Nakashima et al. 2008; Honma et al. 2002), chondrogenesis (Shen et al. 1997), cell growth, cell differentiation (Sun & Taneja 2000), immune response, and apoptosis (Li et al. 2002). Our enrichment analysis of Hepg2 co-factors is consistent with the link between Bhlhe40 expression and hepatic clock and metabolic functions of the liver (Shen et al. 2014; Noshiro et al. 2009). Gm12878 co-factors are enriched for cell differentiation, and signaling pathway which are related to inhibition of cell growth and immune response. Enrichment of with BMP response in leukemia cell line is consistent with stimulation of BMP response in certain kinds of leukemia (Crispino & Le Beau 2012).

2. **CEBPB** - The enrichment analysis of Gm12878 co-factors supports the known roles of CEBPB in the “regulation of genes involved in immune and inflammatory responses” (Chinery et al. 1997), “binding to the IL-1 response element in the IL-6 gene, as well as to regulatory regions of several acute-phase and cytokine genes” (Akira et al. 1990), high induction of CEBPB in blood leukocytes to strengthen muscle (Harries et al. 2012) etc. Association of CEBPB in AML (Acute myeloid leukemia) (Gery et al. 2005) is known, where encouragingly coagulation is enriched among the co-factor functions. Studies have found metastasis in Helas3 via ER stress of unfolded protein response (Brem et al. 2013; Mujcic et al. 2013) and GO analysis shows that the Helas3 co-factors are enriched for ER and unfolded protein response, strongly supporting CEBPB’s role. The function of liver and lung depend of the circadian cycle (Sukumaran et al. 2011; Vollmers et al. 2009).
3. **EP300** - EP300 is acetyl-transferase gene involved in tumor suppression (Campbell et al. 2004; Gayther et al. 2000), cell proliferation specially in myeloproliferative disorders (Steensma et al. 2006), enhance beta-catenin activity (Kimbrel & Kung 2009), chromatin modelling (Campbell et al. 2004), alu-expression (Dieci et al. 2013), induction of epithelial and mesenchymal proteins and cell-adhesion (Krubasik et al. 2006) etc. These are broadly consistent with the enrichment analysis. Enrichment of cell signaling, cell communication in epithelial cancer, limb bud formation in H1hesc, different type of immune and cellular response in normal blood and liver cancer etc. Co-factors identified in Sknsh (brain cancer) are enriched for cortex related hormone-secretion and stimulus, drug response etc. In literature also, there are many evidence about involvement of EP3oo with neuronal disease and its potential as drug for neuronal disorders (Salisbury et al. 1998; Drake Jr. et al. 1993; Kutcher et al. 1987; Kalayam & Alexopoulos 1999). Not surprisingly, EP300 co-factors in liver are involved in response to alcohol and several other metabolic processes. Interestingly, enrichment of several hormone-mediated processes is consistent with the role of EP300 in hepatic encephalopathy (Kügler et al. 1992).
4. **FOS** - FOS processes many extracellular signals via NOTCH signaling (Henken et al. 2012), or stimulating transcription of AP-1 responsive genes (Chiu et al. 1988). Therefore, it is not surprising to see enrichment of various type cell-signaling terms among FOS co-factors. FOS is also involved in other cellular events like differentiation and survival, hypoxia and EMT (epithelial-mesenchymal-transition) (Tulchinsky 2000), metastatis growth in mammary epithelial cells (Langer et al. 2006; Fialka et al. 1996). Furthermore, FOS is a predictor for decreased survival rate in breast cancer (Bland et al. 1995) and is induced by VEGF which plays an important role in the neovascularization in primary breast cancer (Hoeben et al. 2004). We found that breast-specific co-factors are enriched for organ regeneration.
5. **GABPA** - GABPA is known for maintaining homeostasis (Giguere 2008), mitochondrial respiration (Yang et al. 2014), and cellular oxidative stress (Yueh & Tukey 2007; Zhang 2006; Nguyen et al. 2003). The enrichment analysis revealed homeostasis in Gm12878, oxygen-containing compound in Hepg2, and DNA replication in H1hesc, which are consistent with literature. In addition, ETS TF family pays role in the development of vasculature in endothelial cell and its progenitor (Sato 2001) and we find similar evidence of the role of GABPA, an Ets-family member, in K562.
6. **JUN** - AP-1 (JUN/FOS) complex modulates apoptosis in blood cells (Liebermann et al. 1998), controls cell proliferation, cell cycle progression (Behrens et al. 1999; Wisdom et al. 1999), and is involved in angiogenesis (Vleugel et al. 2006; Vasilevskaya & O’Dwyer 1999).

This gene is the putative transforming gene of avian sarcoma virus. We find enrichment of defense mechanism, immune response, homeostasis, estrogen response etc. in blood cells (Gm12878, Huvec, K562). There is also some evidence of involvement of JUN in the development of liver tumor (Eferl et al. 2003), and cervical cancer (Naumann et al. 1998) via co-factors.

7. **MAFK** - MAFK regulates globin genes and plays significant role in coagulation system during embryonic growth and placental development (Isermann & Nawroth 2006). In addition to that, perturbation in MAFK function is highly associated with carcinogenesis, especially leukemia (Kannan et al. 2012; Igarashi et al. 1995; Shyu et al. 2005; Hwang et al. 2013; Lu et al. 1994). Consistently, we found enrichment of mitotic cell cycle in stem cells, meiosis in liver cancer cell, and response to various metal ion in K562 cell.
8. **MAZ** - MAZ regulates MMP genes, gamma fibrinozen, and serum amyliod A (Ray et al. 2003; Ray et al. 2004) which are consistent with the enriched terms among K562 co-factors, blood coagulation, and hemostasis. Immune and viral response functions among Gm12878 co-factors are supported by the study that MAZ plays functional role in CD4 expression (Duncan et al. 1995).
9. **MYC** - MYC is an oncogene, and in Huvec and Gm12878, we found enrichment of cell cycle check points, DNA damage consistent with its role as oncogene. We found that MYC co-factors of H1hesc are enriched for spinal cord development, glial cell fat regulation, limb bud formation, consistent with its role in determining growth size (Zhong et al. 2006), controlling glial cell in stem cells (Kim et al. 2008; Yokoyama et al. 2014), developing limb link with skeletal size (Ota et al. 2007). Gonadotropin up-regulates myeloid protein leukemia-1 (Chen et al. 2010) and is induced by MYC expression (Delidow et al. 1990). In addition to that, MYC regulates intestinal intraepithelial lymphocytes and is involved in the homeostasis of adult intestinal epithelium (Bettess et al. 2005; Jiang et al. 2010). Consistently, K562 co-factors show enrichment of gonadotropin protein, intestinal epithelial cell differentiation along with cell cycle and cell-cell signaling. MCF co-factors were found to be enriched for viral transcription. It has been shown that knockdown of MYC inhibits breast tumor growth by RNA interference which treats cancer by viral infection (Milner 2003; Wang et al. 2005) which is consistent with our findings that breats tumor (Mcf7) cell co-factors are enriched with viral transcription and high RNA production by carbon catabolite.
10. **NRF1** - NRF1 activates the expression of key metabolic genes regulating cellular growth and nuclear genes required for mitochondrial respiration (Li et al. 1999; Evans & Scarpulla 1989; Choi et al. 2004; Vercauteren et al. 2006). We found an enrichment of terms related to mitochondrial respiration and biosynthetic process in most cell lines. Interestingly, the enrichment is evident even though cell specific co-factors are used. Interestingly, in K562 NRF1 shows enrichment of a diverse set of terms, several of which are consistent with literature, e.g. association with neurite outgrowth in rodent (Chang et al. 2005), oxidative stress response (Biswas & Chan 2010), mitochondrial biogenesis (Carew et al. 2004) etc.
11. **REST** - REST acts as a repressor neuronal genes in non-neuronal cell types, and has activation role in neuronal functions (Kuwabara et al. 2004; Huang et al. 1999; Naruse et al. 1999). In the enrichment analysis of REST co-factors, gliblastoma cell line shows cognition, memory, pattern recognition etc. Furthermore, Pancreas cell line shows enrichment for cell differentiation which is consistent with the role of NRSF/REST in pancreas via induction of

Pax4 gene (Kemp et al. 2003). We see extremely high enrichment (>93) of intestinal epithelial cell differentiation among K562 co-factors. We did not find any direct support for this, however, there is evidence that lung and colon epithelial cells show abnormal expression of NRSF in respective cancers (Su et al. 2006).

12. **RFX5** - A lack of MHC-II expression results in a severe immunodeficiency syndrome called MHC-II deficiency, or the bare lymphocyte syndrome (Villard et al. 2000). Helas3 co-factors are enriched for immune response related terms. RFX5 regulates collagen gene expression (Xu et al. 2003), which in turn modulate angiogenesis (Twardowski et al. 2007). Consistently, K562 co-factors are enriched with positive regulation of angiogenesis. RFX5 is found to be up-regulated in primary lung budding and mesenchymal cells of branchial arches and stomach in sub-epithelial layer of mouse (Millien et al. 2008). Consistently, Hepg2 co-factors are enriched with epithelial tube branching involved in lung morphogenesis. RFX5 complex interacts with the collagen in human fibroblasts (Sengupta et al. 2002) and consistently, regulation of fibroblast proliferation is enriched in Gm12878.
13. **USF1** – Upstream regulatory factor 1 is known for regulating multiple genes of glucose and lipid metabolism (Naukkarinen et al. 2005; Pajukanta et al. 2004). In almost all cell lines, the USF1 co-factors are enriched with hormone mediated signaling pathway; especially epithelium cancer cell line (A549) shows enrichment of lipid metabolism related terms. In addition, ovulation, reproductive process, female pregnancy are significantly enriched in A549 co-factors which is consistent with suppression of Follicle-stimulating hormone receptor activity by USF1 (Putowski et al. 2004).
14. **YY1** - YY1 is known as ubiquitous TF. Still the co-factor enrichment analysis shows some of its cell-specific roles. For example, enrichment of epithelial cell maturation in prostate gland development in Gm12878 co-factors (Kashyap & Bonavida 2014), various biosynthetic process in K562 and Hct116 (Bennett et al. 1999; Villagra et al. 2007) are consistent with literature. Other ontology associated with YY1 co-factors are cell cycle/DNA damage (Rizkallah & Hurt 2009; Affar et al. 2006). Nt2d1 and Hepg2 co-factors are enriched with cell cycles and DNA metabolic process respectively.
15. **ZNF143** – ZNF143 plays role as one of the key components of three-dimensional chromatin structure (Bailey et al. 2015; Heidari et al. 2014), regulates dna-replication and cell-cycle-associated genes (Lu et al. 2012; Izumi et al. 2010). Among the co-factors, we found enrichment of cell-adhesion and cell-proliferation in 3 out of 4 cells.

Supplemental Note 8. We have detailed lists of cell type-specific co-factors for all TFs that we can provide as supplementary online material, and which can serve as a resource for others. Here we discuss a few cases which demonstrate that the co-factors revealed by *TRISECT* are supported by previous experimental research. Recall that P300 is not a typical TF with a DNA binding motif. It nevertheless is expected to interact with other DNA-binding co-factors to achieve specificity. *TRISECT* revealed TEAD as one of the most influential co-factors of P300 in multiple cell types. Indeed, TEAD is known to form a complex with P300 providing locus-specificity to P300 (Fujii et al. 2012). Likewise, CEBP is known to recruit P300 (Schwartz et al. 2003) and ATF interacts with P300's HAT domain (Karanam et al. 2007), and both were detected as P300's ubiquitous co-factors. NR2F2 (also known as COUP-TF2) has a liver-specific function (Zhang et al. 2002) and is known to interact with P300, although this was shown in a different context (Bailey et al. 1998). It is this interesting that our method detects NR2F2 as Hepg2-specific co-factor of P300. Likewise, members of GATA family are core

regulators in liver. We found that in many (but not all) cell lines, and notably in HepG2, members of GATA family are co-factors of P300, consistent with (Dai & Markham 2001). Serum response factor (SRF) is a ubiquitous protein and with a specific function in liver (Sun et al. 2009). FOXA TFs are critical for liver development and function (Lee et al. 2005). Our analysis reveals FOXA TFs as HepG2-specific co-factors of SRF. On the other hand, we found ELK4 to be broadly used co-factor of SRF, consistent with their broad expression and known physical interaction with SRF (Shore & Sharrocks 1994). As yet another example, PAX1 and SOX4 are key TFs in embryogenesis, and both are revealed as co-factors of the core promoter factor *Tata Binding Protein* (TBP) specifically in hESC.

References

Affar, E.B. et al., 2006. Essential dosage-dependent functions of the transcription factor yin yang 1 in late embryonic development and cell cycle progression. *Molecular and cellular biology*, 26(9), pp.3565–3581.

Akira, S. et al., 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *The EMBO journal*, 9(6), pp.1897–1906.

Bailey, P. et al., 1998. The orphan nuclear receptor, COUP-TF II, inhibits myogenesis by post-transcriptional regulation of MyoD function: COUP-TF II directly interacts with p300 and myoD. *Nucleic acids research*, 26(23), pp.5501–5510.

Bailey, S.D. et al., 2015. ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. *Nature communications*, 2, p.6186.

Behrens, A., Sibilia, M. & Wagner, E.F., 1999. Amino-terminal phosphorylation of c-Jun regulates stress-induced apoptosis and cellular proliferation. *Nature genetics*, 21(3), pp.326–329.

Bennett, M.K. et al., 1999. Co-stimulation of promoter for low density lipoprotein receptor gene by sterol regulatory element-binding protein and Sp1 is specifically disrupted by the yin yang 1 protein. *The Journal of biological chemistry*, 274(19), pp.13025–13032.

Bettess, M.D. et al., 2005. c-Myc is required for the formation of intestinal crypts but dispensable for homeostasis of the adult intestinal epithelium. *Molecular and cellular biology*, 25(17), pp.7868–7878.

Biswas, M. & Chan, J.Y., 2010. Role of Nrf1 in antioxidant response element-mediated gene expression and beyond. *Toxicology and Applied Pharmacology*, 244(1), pp.16–20.

Bland, K.I. et al., 1995. Oncogene protein co-expression. Value of Ha-ras, c-myc, c-fos, and p53 as prognostic discriminants for breast carcinoma. *Annals of surgery*, 221(6), pp.706–720.

Brem, G.J., Mylonas, I. & Brüning, A., 2013. Eeyarestatin causes cervical cancer cell sensitization to bortezomib treatment by augmenting ER stress and CHOP expression. *Gynecologic Oncology*, 128(2), pp.383–390.

Campbell, I.G., Choong, D. & Chenevix-Trench, G., 2004. No germline mutations in the histone acetyltransferase gene EP300 in BRCA1 and BRCA2 negative families with breast cancer and gastric, pancreatic, or colorectal cancer. *Breast cancer research : BCR*, 6(4), pp.R366–71.

Carew, J.S. et al., 2004. Increased mitochondrial biogenesis in primary leukemia cells: the role of endogenous nitric oxide and impact on sensitivity to fludarabine. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K*, 18(12), pp.1934–1940.

Chang, W.-T. et al., 2005. A novel function of transcription factor alpha-Pal/NRF-1: increasing neurite outgrowth. *Biochemical and biophysical research communications*, 334(1), pp.199–206.

Chen, S.U. et al., 2010. Human chorionic gonadotropin up-regulates expression of myeloid cell leukemia-1 protein in human granulosa-lutein cells: Implication of corpus luteum rescue and ovarian hyperstimulation syndrome. *Journal of Clinical Endocrinology and Metabolism*, 95(8), pp.3982–3992.

Chinery, R. et al., 1997. Antioxidant-induced nuclear translocation of CCAAT/enhancer-binding protein beta. A critical role for protein kinase A-mediated phosphorylation of Ser299. *The Journal of biological chemistry*, 272(48), pp.30356–30361.

Chiu, R. et al., 1988. The c-Fos protein interacts with c-Jun/AP-1 to stimulate transcription of AP-1 responsive genes. *Cell*, 54(4), pp.541–552.

Choi, Y.S. et al., 2004. In vitro methylation of nuclear respiratory factor-1 binding site suppresses the promoter activity of mitochondrial transcription factor A. *Biochemical and Biophysical Research Communications*, 314(1), pp.118–122.

Crispino, J.D. & Le Beau, M.M., 2012. BMP Meets AML: Induction of BMP Signaling by a Novel Fusion Gene Promotes Pediatric Acute Leukemia. *Cancer Cell*, 22(5), pp.567–568.

Dai, Y.S. & Markham, B.E., 2001. p300 Functions as a Coactivator of Transcription Factor GATA-4. *Journal of Biological Chemistry*, 276(40), pp.37178–37185.

Delidow, B.C., White, B.A. & Peluso, J.J., 1990. Gonadotropin induction of c-fos and c-myc expression and deoxyribonucleic acid synthesis in rat granulosa cells. *Endocrinology*, 126(5), pp.2302–2306.

Dieci, G. et al., 2013. Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1829(3-4), pp.296–305.

Drake Jr., M.E. et al., 1993. Neuropsychological and event-related potential correlates of nonepileptic seizures. *J.Neuropsychiatry Clin.Neurosci.*, 5(0895-0172 SB - IM), pp.102–104.

Duda, R., Hart, P. & Stork, D., 2001. Pattern Classification. *New York: John Wiley, Section*, p.680.

Duncan, D.D. et al., 1995. A Myc-associated zinc finger protein binding site is one of four important functional regions in the CD4 promoter. *Molecular and cellular biology*, 15(6), pp.3179–3186.

Eferl, R. et al., 2003. Liver tumor development. c-Jun antagonizes the proapoptotic activity of p53. *Cell*, 112(2), pp.181–192.

Evans, M.J. & Scarpulla, R.C., 1989. Interaction of nuclear factors with multiple sites in the

somatic cytochrome c promoter. Characterization of upstream NRF-1, ATF, and intron Sp1 recognition sequences. *Journal of Biological Chemistry*, 264(24), pp.14361–14368.

Fialka, I. et al., 1996. The estrogen-dependent C-junER protein causes a reversible loss of mammary epithelial cell polarity involving a destabilization of adherens junctions. *Journal of Cell Biology*, 132(6), pp.1115–1132.

Fujii, M. et al., 2012. Convergent signaling in the regulation of connective tissue growth factor in malignant mesothelioma: TGFbeta signaling and defects in the Hippo signaling cascade. *Cell cycle (Georgetown, Tex.)*, 11(18), pp.3373–3379.

Gayther, S.A. et al., 2000. Mutations truncating the EP300 acetylase in human cancers. *Nature genetics*, 24(3), pp.300–303.

Gery, S. et al., 2005. Transcription profiling of C/EBP targets identifies Per2 as a gene implicated in myeloid leukemia. *Blood*, 106(8), pp.2827–2836.

Giguere, V., 2008. Transcriptional control of energy homeostasis by the estrogen-related receptors. *Endocrine reviews*, 29(6), pp.677–696. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18664618.

Harries, L.W. et al., 2012. CCAAT-enhancer-binding protein-beta expression in vivo is associated with muscle strength. *Aging Cell*, 11(2), pp.262–268.

Heidari, N. et al., 2014. Genome-wide map of regulatory interactions in the human genome. *Genome research*, 24(12), pp.1905–1917.

Henken, F.E. et al., 2012. The functional role of Notch signaling in HPV-mediated transformation is dose-dependent and linked to AP-1 alterations. *Cellular oncology (Dordrecht)*, 35(2), pp.77–84.

Hoeben, A. et al., 2004. Vascular endothelial growth factor and angiogenesis. *Pharmacological reviews*, 56(4), pp.549–580.

Honma, S. et al., 2002. Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature*, 419(6909), pp.841–844.

Huang, Y., Myers, S.J. & Dingledine, R., 1999. Transcriptional repression by REST: recruitment of Sin3A and histone deacetylase to neuronal genes. *Nature neuroscience*, 2(10), pp.867–872.

Hwang, Y.-J. et al., 2013. MafK positively regulates NF-κB activity by enhancing CBP-mediated p65 acetylation. *Scientific reports*, 3, p.3242. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3832860&tool=pmcentrez&rendertype=abstract.

Igarashi, K. et al., 1995. Conditional expression of the ubiquitous transcription factor MafK induces erythroleukemia cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 92(16), pp.7445–7449.

Isermann, B. & Nawroth, P.P., 2006. The role of platelets during reproduction. *Pathophysiology of haemostasis and thrombosis*, 35(1-2), pp.23–27.

Izumi, H. et al., 2010. Role of ZNF143 in tumor growth through transcriptional regulation of DNA replication and cell-cycle-associated genes. *Cancer Science*, 101(12), pp.2538–

2545.

Jain, A.K. & Dubes, R.C., 1988. *Algorithms for Clustering Data*,

Jiang, W. et al., 2010. c-Myc controls the development of CD8alphaalpha TCRalphabeta intestinal intraepithelial lymphocytes from thymic precursors by regulating IL-15-dependent survival. *Blood*, 115(22), pp.4431–4438.

Kalayam, B. & Alexopoulos, G.S., 1999. Prefrontal dysfunction and treatment response in geriatric depression. *Archives of general psychiatry*, 56(8), pp.713–718.

Kannan, M.B., Solovieva, V. & Blank, V., 2012. The small MAF transcription factors MAFF, MAFG and MAFK: Current knowledge and perspectives. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1823(10), pp.1841–1846.

Karanam, B. et al., 2007. Multiple roles for acetylation in the interaction of p300 HAT with ATF-2. *Biochemistry*, 46(28), pp.8207–8216.

Kashyap, V. & Bonavida, B., 2014. Role of YY1 in the pathogenesis of prostate cancer and correlation with bioinformatic data sets of gene expression. *Genes & cancer*, 5(3-4), pp.71–83.

Kemp, D.M., Lin, J.C. & Habener, J.F., 2003. Regulation of Pax4 paired homeodomain gene by neuron-restrictive silencer factor. *Journal of Biological Chemistry*, 278(37), pp.35057–35062.

Kim, H. et al., 2008. Notch-regulated oligodendrocyte specification from radial glia in the spinal cord of zebrafish embryos. *Developmental Dynamics*, 237(8), pp.2081–2089.

Kimbrel, E.A. & Kung, A.L., 2009. The F-box protein beta-TrCp1/Fbw1a interacts with p300 to enhance beta-catenin transcriptional activity. *The Journal of biological chemistry*, 284(19), pp.13033–13044.

Krubasik, D. et al., 2006. Absence of p300 induces cellular phenotypic changes characteristic of epithelial to mesenchyme transition. *British journal of cancer*, 94(9), pp.1326–1332.

Kügler, C.F. et al., 1992. Visual event-related P300 potentials in early portosystemic encephalopathy. *Gastroenterology*, 103(1), pp.302–310.

Kutcher, S.P. et al., 1987. Auditory P300 in borderline personality disorder and schizophrenia. *Archives of general psychiatry*, 44(7), pp.645–650.

Kuwabara, T. et al., 2004. A small modulatory dsRNA specifies the fate of adult neural stem cells. *Cell*, 116(6), pp.779–793.

Langer, S. et al., 2006. Jun and Fos family protein expression in human breast cancer: Correlation of protein expression and clinicopathological parameters. *European Journal of Gynaecological Oncology*, 27(4), pp.345–352.

Lee, C.S. et al., 2005. The initiation of liver development is dependent on Foxa transcription factors. *Nature*, 435(7044), pp.944–947. Available at: http://dx.doi.org/10.1038/nature03649.

Lenhard, B. & Wasserman, W.W., 2002. TFBS: Computational framework for transcription factor binding site analysis. *Bioinformatics (Oxford, England)*, 18(8), pp.1135–6. Available

at: http://www.ncbi.nlm.nih.gov/pubmed/12176838.

Li, B., Holloszy, J.O. & Semenkovich, C.F., 1999. Respiratory uncoupling induces ??-aminolevulinate synthase expression through a nuclear respiratory factor-1-dependent mechanism in HeLa cells. *Journal of Biological Chemistry*, 274(25), pp.17534–17540.

Li, Y. et al., 2002. Abundant expression of Dec1/stra13/sharp2 in colon carcinoma: its antagonizing role in serum deprivation-induced apoptosis and selective inhibition of procaspase activation. *The Biochemical journal*, 367(Pt 2), pp.413–422.

Liebermann, D.A., Gregory, B. & Huffman, B., 1998. AP-1 (FOS/JUN) transcription factors in hematopoietic differentiation and apoptosis (Review). *International Journal of Oncology*, 12(3), pp.685–700.

Lu, S.J. et al., 1994. Retroviral integration within the Fli-2 locus results in inactivation of the erythroid transcription factor NF-E2 in Friend erythroleukemias: evidence that NF-E2 is essential for globin expression. *Proceedings of the National Academy of Sciences of the United States of America*, 91(18), pp.8398–8402.

Lu, W. et al., 2012. ZNF143 transcription factor mediates cell survival through upregulation of the GPX1 activity in the mitochondrial respiratory dysfunction. *Cell death & disease*, 3, p.e422. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3542592&tool=pmcentrez&rendertype=abstract.

Millien, G. et al., 2008. Characterization of the mid-foregut transcriptome identifies genes regulated during lung bud induction. *Gene Expression Patterns*, 8(2), pp.124–139.

Milner, J., 2003. RNA interference for treating cancers caused by viral infection. *Expert opinion on biological therapy*, 3(3), pp.459–467.

Mujcic, H. et al., 2013. Hypoxic activation of the PERK/eIF2α arm of the unfolded protein response promotes metastasis through induction of LAMP3. *Clinical Cancer Research*, 19(22), pp.6126–6137.

Nakashima, A. et al., 2008. DEC1 modulates the circadian phase of clock gene expression. *Molecular and cellular biology*, 28(12), pp.4080–4092.

Naruse, Y. et al., 1999. Neural restrictive silencer factor recruits mSin3 and histone deacetylase complex to repress neuron-specific target genes. *Proceedings of the National Academy of Sciences of the United States of America*, 96(24), pp.13691–13696.

Naukkarinen, J. et al., 2005. USF1 and dyslipidemias: Converging evidence for a functional intronic variant. *Human Molecular Genetics*, 14(17), pp.2595–2605.

Naumann, M. et al., 1998. Coordinate activation of activator protein 1 and inflammatory cytokines in response to Neisseria gonorrhoeae epithelial cell contact involves stress response kinases. *The Journal of experimental medicine*, 188(7), pp.1277–1286.

Nguyen, T., Sherratt, P.J. & Pickett, C.B., 2003. Regulatory mechanisms controlling gene expression mediated by the antioxidant response element. *Annual review of pharmacology and toxicology*, 43, pp.233–260.

Noshiro, M. et al., 2009. Liver X receptors (LXRalpha and LXRbeta) are potent regulators for

hepatic Dec1 expression. *Genes to cells : devoted to molecular & cellular mechanisms*, 14(1), pp.29–40.

Ota, S. et al., 2007. Activities of N-Myc in the developing limb link control of skeletal size with digit separation. *Development (Cambridge, England)*, 134(8), pp.1583–1592.

Pajukanta, P. et al., 2004. Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nature genetics*, 36(4), pp.371–376.

Putowski, L.T. et al., 2004. Human follicle-stimulating hormone receptor (FSH-R) promoter/enhancer activity is inhibited by transcriptional factors, from the upstream stimulating factors family, via E-box and newly identified initiator element (Inr) in FSH-R non-expressing cells. *Gynecological endocrinology : the official journal of the International Society of Gynecological Endocrinology*, 19(1), pp.9–17.

Ray, A. et al., 2003. Induction of matrix metalloproteinase 1 gene expression is regulated by inflammation-responsive transcription factor SAF-1 in osteoarthritis. *Arthritis and Rheumatism*, 48(1), pp.134–145.

Ray, A. et al., 2004. Overexpression of serum amyloid A-activating factor 1 inhibits cell proliferation by the induction of cyclin-dependent protein kinase inhibitor p21WAF-1/Cip-1/Sdi-1 expression. *Journal of immunology (Baltimore, Md. : 1950)*, 172(8), pp.5006–5015.

Rizkallah, R. & Hurt, M.M., 2009. Regulation of the transcription factor YY1 in mitosis through phosphorylation of its DNA-binding domain. *Molecular biology of the cell*, 20(22), pp.4766–4776.

Salisbury, D.F. et al., 1998. First-episode schizophrenic psychosis differs from first-episode affective psychosis and controls in P300 amplitude over left temporal lobe. *Archives of general psychiatry*, 55(2), pp.173–180.

Sato, Y., 2001. Role of ETS family transcription factors in vascular development and angiogenesis. *Cell structure and function*, 26(1), pp.19–24.

Schwartz, C. et al., 2003. Recruitment of p300 by C/EBP?? triggers phosphorylation of p300 and modulates coactivator activity. *EMBO Journal*, 22(4), pp.882–892.

Sengupta, P.K., Fargo, J. & Smith, B.D., 2002. The RFX family interacts at the collagen (COL1A2) start site and represses transcription. *Journal of Biological Chemistry*, 277(28), pp.24926–24937.

Shen, L. et al., 2014. Hepatic differentiated embryo-chondrocyte-expressed gene 1 (Dec1) inhibits sterol regulatory element-binding protein-1c (Srebp-1c) expression and alleviates fatty liver phenotype. *The Journal of biological chemistry*, 289(34), pp.23332–23342.

Shen, M. et al., 1997. Molecular characterization of the novel basic helix-loop-helix protein DEC1 expressed in differentiated human embryo chondrocytes. *Biochemical and biophysical research communications*, 236(2), pp.294–298.

Shore, P. & Sharrocks, A.D., 1994. The transcription factors Elk-1 and serum response factor interact by direct protein-protein contacts mediated by a short region of Elk-1. *Molecular and cellular biology*, 14(5), pp.3283–3291.

Shyu, Y.-C. et al., 2005. Sumoylation of p45/NF-E2: nuclear positioning and transcriptional

584 activation of the mammalian beta-like globin gene locus. *Molecular and cellular biology*,
585 25(23), pp.10365–10378.

586 Steensma, D.P. et al., 2006. More on Myb in myelofibrosis: molecular analyses of MYB and
587 EP300 in 55 patients with myeloproliferative disorders. *Blood*, 107(4), pp.1733–5; author
588 reply 1735.

589 Su, X. et al., 2006. Abnormal expression of REST/NRSF and Myc in neural stem/progenitor
590 cells causes cerebellar tumors by blocking neuronal differentiation. *Molecular and cellular*
591 *biology*, 26(5), pp.1666–1678.

592 Sukumaran, S. et al., 2011. Light-dark oscillations in the lung transcriptome: implications for
593 lung homeostasis, repair, metabolism, disease, and drug action. *Journal of applied*
594 *physiology (Bethesda, Md.: 1985)*, 110(6), pp.1732–1747.

595 Sun, H. & Taneja, R., 2000. Stra13 expression is associated with growth arrest and represses
596 transcription through histone deacetylase (HDAC)-dependent and HDAC-independent
597 mechanisms. *Proceedings of the National Academy of Sciences of the United States of*
598 *America*, 97(8), pp.4058–4063.

599 Sun, K. et al., 2009. Hepatocyte expression of serum response factor is essential for liver
600 function, hepatocyte proliferation and survival, and postnatal body growth in mice.
601 *Hepatology (Baltimore, Md.)*, 49(5), pp.1645–54. Available at:
602 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2810404&tool=pmcentrez&ren
603 dertype=abstract.

604 Tulchinsky, E., 2000. Fos family members: Regulation, structure and role in oncogenic
605 transformation. *Histology and Histopathology*, 15(3), pp.921–928.

606 Twardowski, T. et al., 2007. Type I collagen and collagen mimetics as angiogenesis promoting
607 superpolymers. *Current pharmaceutical design*, 13(35), pp.3608–3621.

608 Vasilevskaya, I.A. & O'Dwyer, P.J., 1999. Effects of geldanamycin on signaling through
609 activator-protein 1 in hypoxic HT29 human colon adenocarcinoma cells. *Cancer*
610 *Research*, 59(16), pp.3935–3940.

611 Vercauteren, K. et al., 2006. PGC-1-related coactivator: immediate early expression and
612 characterization of a CREB/NRF-1 binding domain associated with cytochrome c
613 promoter occupancy and respiratory growth. *Molecular and cellular biology*, 26(20),
614 pp.7409–7419.

615 Villagra, A. et al., 2007. Histone deacetylase 3 down-regulates cholesterol synthesis through
616 repression of lanosterol synthase gene expression. *Journal of Biological Chemistry*,
617 282(49), pp.35457–35470.

618 Villard, J. et al., 2000. A functionally essential domain of RFX5 mediates activation of major
619 histocompatibility complex class II promoters by promoting cooperative binding between
620 RFX and NF-Y. *Molecular and cellular biology*, 20(10), pp.3364–3376.

621 Vleugel, M.M. et al., 2006. c-Jun activation is associated with proliferation and angiogenesis in
622 invasive breast cancer. *Human Pathology*, 37(6), pp.668–674.

623 Vollmers, C. et al., 2009. Time of feeding and the intrinsic circadian clock drive rhythms in
624 hepatic gene expression. *Proceedings of the National Academy of Sciences of the United*

States of America, 106(50), pp.21453–21458.

Wang, Y. et al., 2005. Knockdown of c-Myc expression by RNAi inhibits MCF-7 breast tumor cells growth in vitro and in vivo. *Breast cancer research : BCR*, 7(2), pp.R220–R228.

Wien, T.U., 2002. an Examination of Indexes for Determining. , 67(3).

Wisdom, R., Johnson, R.S. & Moore, C., 1999. c-Jun regulates cell cycle progression and apoptosis by distinct mechanisms. *The EMBO journal*, 18(1), pp.188–197.

Worsley Hunt, R. & Wasserman, W.W., 2014. *Non-targeted transcription factors motifs are a systemic component of ChIP-seq datasets.*, Available at: http://genomebiology.com/2014/15/7/412/abstract.

Xu, Y. et al., 2003. Interferon gamma repression of collagen (COL1A2) transcription is mediated by the RFX5 complex. *The Journal of biological chemistry*, 278(49), pp.49134–49144.

Yang, Z.-F. et al., 2014. GABP transcription factor (nuclear respiratory factor 2) is required for mitochondrial biogenesis. *Molecular and cellular biology*, 34(17), pp.3194–3201.

Yokoyama, K.D., Zhang, Y. & Ma, J., 2014. Tracing the Evolution of Lineage-Specific Transcription Factor Binding Sites in a Birth-Death Framework. *PLoS computational biology*, 10(8), p.e1003771. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25144359.

Yueh, M.F. & Tukey, R.H., 2007. Nrf2-Keap1 Signaling pathway regulates human UGT1A1 expression in vitro and in transgenic UGT1 mice. *Journal of Biological Chemistry*, 282(12), pp.8749–8758.

Zhang, D.D., 2006. Mechanistic studies of the Nrf2-Keap1 signaling pathway. *Drug metabolism reviews*, 38(4), pp.769–789.

Zhang, P. et al., 2002. Expression of COUP-TFII in metabolic tissues during development. *Mechanisms of Development*, 119(1), pp.109–114.

Zhong, W. et al., 2006. Hypertrophic growth in cardiac myocytes is mediated by Myc through a Cyclin D2-dependent pathway. *The EMBO journal*, 25(16), pp.3869–3879.