**Aim 1: To determine the prevalence of transient Hit-and-Run interactions in GRNs**

***Rationale***:Our experiments using the cell-based *TARGET* assay have provided evidence for *Hit-and-Run* transcription for two master TFs in N-signaling bZIP1(Para, Li et al. 2014, Doidy, Li et al. 2016) and NLP7 (Aim 2A, Preliminary results, Fig. X) (Alvarez 2019). We now ask the question: How prevalent is the *Hit-and-Run* mechanism across TFs? In Aim 1, we will identify *Hit-and-Run* candidate TFs among members of all 70 TF families of Arabidopsis.

**1A. High-throughput identification of direct TF-regulated and TF-bound targetsin root cells**

***Innovation***: We have implemented several innovations to the *TARGET* assay to scale-up the throughput, as described in Brooks et al. 2019, Nature Comm.)(Brooks, Cirrone et al. 2019) and summarized below. The higher throughput *TARGET* approach allows us to identify direct TF-targets based on TF-regulation and TF-binding for 150 TFs within 6 months-time. This is far faster than if done in transgenic plants and allows us to capture transient TF-target interactions that are missed *in planta*. By integrating TF-regulation and TF-binding datasets (Aim 1B), we will identify candidate *Hit-and-Run* TFs. Specifically, we will identify the TF-regulated and TF-bound targets of 150 TFs using this optimized *TARGET* assay. While a majority of these TFs (99/150) respond transcriptionally to N-treatment in shoots and/or roots, we also include a representative for all 70 TF families in Arabidopsis to broaden the impact of our findings (Palaniswamy, James et al. 2006, Pruneda-Paz, Breton et al. 2014, Yang, Jin et al. 2016).

***Approach***: To determine the genome-wide targets directly regulated by the prioritized 150 TFs, we will use the

*TARGET* system on root cells isolated from Arabidopsis seedlings grown in ½ MS media for 10 days. The combination of +DEX and +CHX treatments will allow us to identify direct TF-targets regulated at 180 min after DEX-induced TF nuclear entry (as described in *Background*). Modifications to the protocol(Brooks, Cirrone et al. 2019) will enable us to screen the 150 TFs in less than 6 months and include: ***i***) an Empty Vector (EV) containing only the GR and no TF, which serves as a control for multiple TFs analyzed on a single day, and ***ii***) pooling of root cells each separately transfected with distinct TFs on vectors containing either RFP or GFP, prior to FACS. These changes enabled us to increase the throughput of *TARGET* up to 24 TF assays/day (e.g. 8 TFs x three replicates)(Brooks, Cirrone et al. 2019), vs. a single TF per day in the original *TARGET* assay (Bargmann 2013). We will also identify targets that are bound by each TF in the same root cell samples at the 180 min time point. To improve the specificity/sensitivity of the standard uChIP protocol (Para et al., 2014; Para 2018) we will use Biotin-ChIP(Sura et al., 2018), which takes advantage of the strong interaction between streptavidin and biotin(Green 1975). By adding the short biotin ligase recognition peptide (BLRP) to the N-terminal of the GR-TF fusion(Fairhead and Howarth 2015) and expressing the BirA enzyme in the same cells, the BLRP-GR-TF fusion protein will be biotinylated, enabling pull down of BLRP-GR-TF-chromatin complex with streptavidin(Kim, Cantor et al. 2009, Fairhead and Howarth 2015). The strong interaction between the streptavidin beads and biotin labeled fusion protein allows stringent washing will enhance the signal to noise ratio (Kim, Cantor et al. 2009).

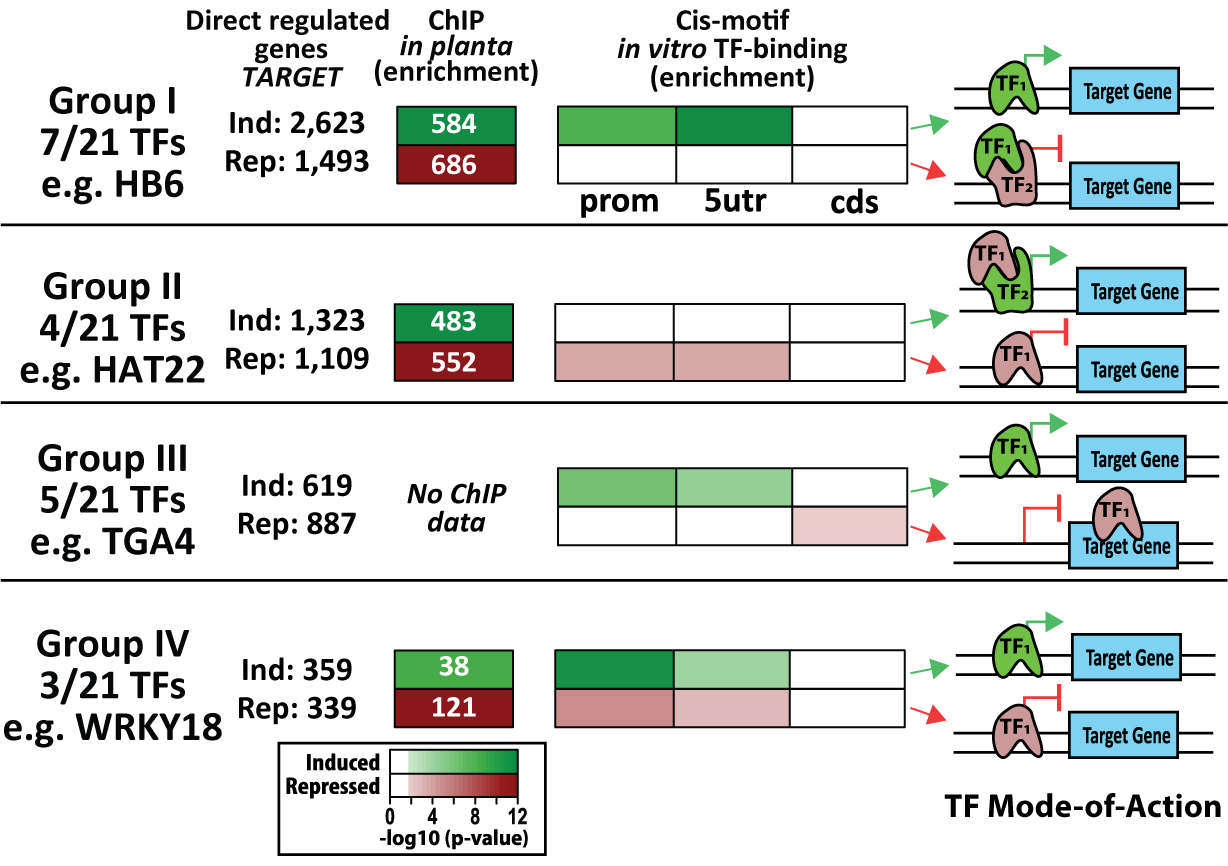


Fig. X -

***Preliminary Results***:Using our enhanced *TARGET* approach, our proof-of-concept study identified direct regulated targets of 33 TFs(Brooks, Cirrone et al. 2019) that were identified as “early” N-responsive genes in roots (Varala, Marshall-Colón et al. 2018). The validated TF-target network generated revealed that 32/33 TFs directly regulated a significant number of the root N-responsive genes. In combination, these TFs regulated 88% (1288/1458) of the root N-responsive genes(Brooks, Cirrone et al. 2019). [MattThis sentence is very unclear. I would consider dropping:]We found that direct TF-regulated targets identified in root cells provide much needed *in vivo* context to *in vitro* identified cis-binding motifs(O'Malley, Huang et al. 2016), i.e. direct TF binding to known cis-motifs could be classified based on the effect it had on gene regulation(Brooks, Cirrone et al. 2019) (Fig. X). We discovered that each TF can act in a dual mode - to induce or repress - depending on the target gene (Fig. X). Our data suggest this dual TF action could operate via direct binding to a cis-element, or indirectly via a partner TF binding (Fig, X). Moreover, a comparison of the direct TF-regulated targets from root cells (*TARGET*) with those bound *in planta,* revealed that for some TFs (HAT22 & HB6) as much as 40-50% of direct TF-regulated targets were TF-bound, whereas among other TFs (bZIP1, NLP7 and WRKY18), only 13-23% of the direct TF-regulated targets were bound (Table. 1)(Brooks, Cirrone et al. 2019). The latter group of TFs includes bZIP1(Para, Li et al. 2014, Doidy, Li et al. 2016) and NLP7 (Preliminary Results, Aim 2A), our two proof-of-concept *Hit-and-Run* TFs.

***Interpretation and Expected outcomes***: In our published study using increased throughput *TARGET*, we demonstrate how this approach can rapidly generate a GRN of validated TF-target edges, serving as a gold-standard to perform Precision/Recall analysis for network inference(Brooks 2019). Additionally, we integrated our cell-based direct TF-target data with such as *in vitro* identified cis-binding motifs enabling us to reveal insights into TF regulatory mechanisms that would have been difficult to discover using only a few TFs (Fig. X). We will now scale our *TARGET* assay and implement the biotin-ChIP approach to identify TF-target interactions, allowing us to rapidly compare the direct TF-regulated and TF-bound targets for the 150 TFs without relying on *in planta* binding data. This will enable us to identify candidate *Hit-and-Run* TFs as ones that have a high proportion of direct TF-regulated targets vs. TF-bound targets, similar to proportions (13-20% overlap) seen for our proof-of concept *Hit-and-Run* TFs, bZIP1 and NLP7 (Table 1, Preliminary Results, Aim 2A). For example, WRKY18 is a candidate *Hit-and-Run* TF with only 23% of direct TF-regulated targets that are also TF-bound (Table 1). Candidate *Hit-and-Run* TFs identified in Aim 1 will be further validated by time-series ChIP and DamID experiments in Aim 2. By analyzing TFs across all 70 TF families, we will be able to address the unresolved question of the prevalence of the *Hit-and-Run* mechanism.

***Potential Problems and Alternative Approaches***: We do not expect significant issues with the biotin-ChIP protocol as it has been used in plants(Sura, Kabza et al. 2017), animals (Kim, Cantor et al. 2009, Matsuda, Mikami et al. 2017) and in cells with a transient expression system (Kulyyassov, Shoaib et al. 2011). However, should problems arise, we can use the previously published uChIP protocol(Para, Li et al. 2018) with the GR-antibody. CHX treatment can also affect gene expression and make it difficult to identify differentially expressed genes. To address this, we will ensure that housekeeping genes behave as expected across experiments and remove genes that are known to have a large and significant response to CHX treatment(Brooks, Cirrone et al. 2019).

**1B. ConnecTF: A database and web tool to integrate and analyze TF-target networks [Dennis thinks that if we run out of space we should vastly shorten this part. It’s useful software but not technically innovative snd distracts from the scientific questions of the prevalence of hit and run and its importance. ]**

***Innovation:*** To analyze the large amount of TF-target data we will generate in Aim 1, we are developing an tool called *ConnecTF*. Most online databases for large sets of TF-targets and cis-elements (e.g. DAP-seq(O'Malley, Huang et al. 2016), CisBP(Weirauch, Yang et al. 2014)), serve only as a repository of genome-wide datasets, and lack the ability to perform meaningful analyses by integrating with other experiments. The *ConnecTF* web tool will enable us - and the research community - to access and analyze the large TF-target regulation/binding datasets we will generate in Aim 1A. An important feature of *ConnecTF* is that it not only enables access to the data, but also provides a user-friendly interface to perform analyses that integrates various types of datasets, including the user’s own genelists and predicted networks. The *ConnecTF* database is capable of handling multiple types of TF-target data such as regulation and binding, generated by us and from other large TF-target studies performed *in vitro* (e.g. DAP-seq(O'Malley, Huang et al. 2016)) and *in vivo*(Song, Huang et al. 2016)*.* The structure of *ConnecTF* database makes it easy to update with new TF-target data and will facilitate our identification of potential *Hit-and-Run* TFs for further analysis (Aim 2), by serving as a platform to integrate the TF-regulation and TF-binding data for 150 TFs from Aim 1A.

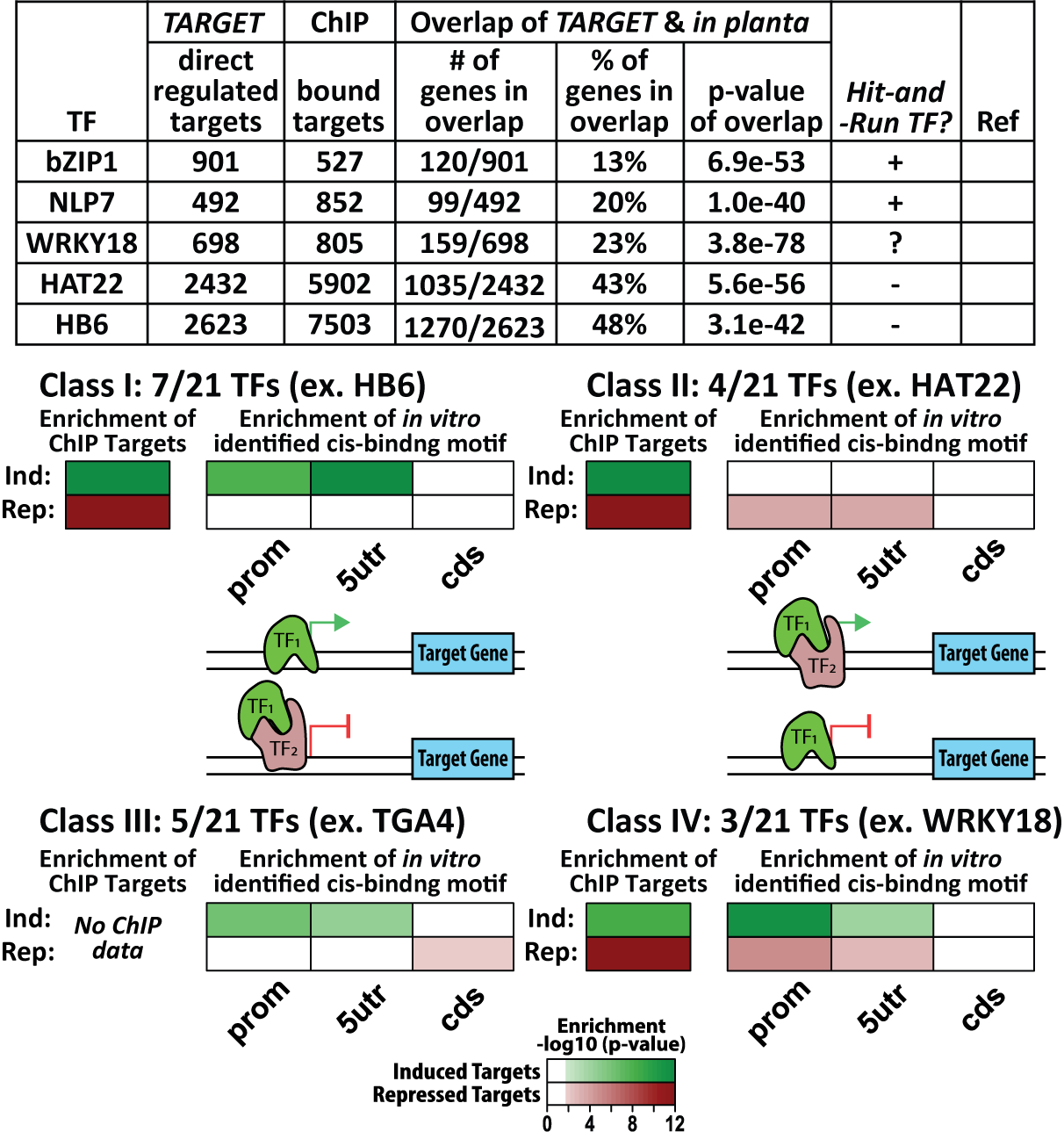


Table 1 -

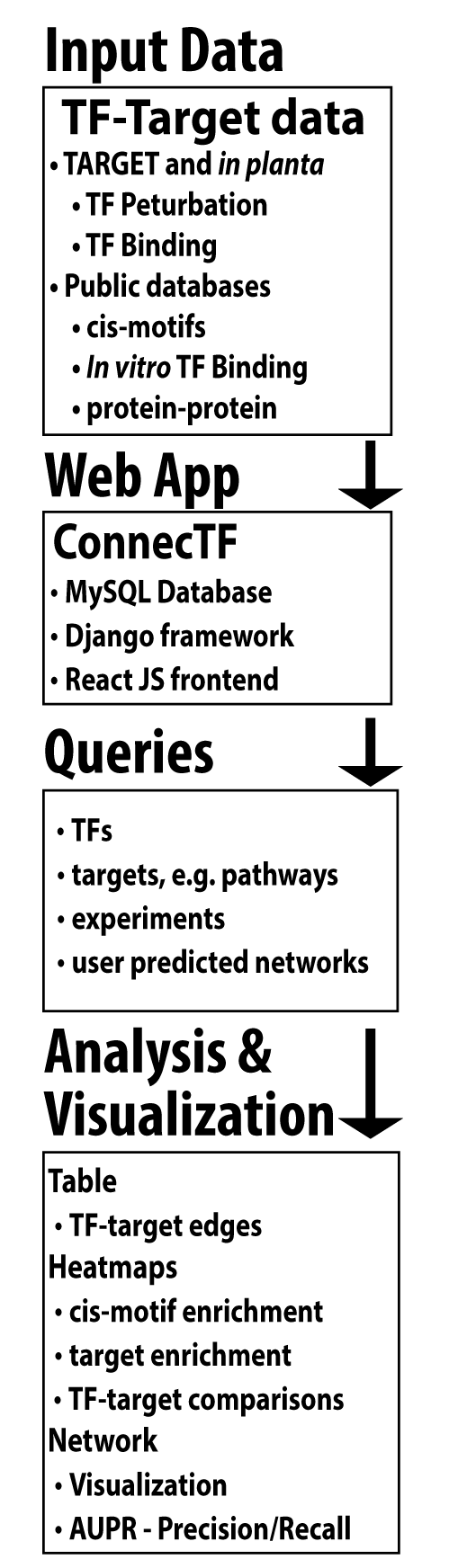
­­

Fig. X -

***Implementation***: The *ConnecTF* database and web tool will serve as a repository of the TF-target data generated in this study (Fig. X, Input). This includes regulated targets identified of a TF in cells (Aims 1A and 2A) and *in planta* (Aim 2B)*,* and TF-bound targets identified using ChIP and DamID (Aim 2A). *ConnecTF* will also include other large-scale datasets such as ChIP data for 21 TFs from Song et al.(Song, Huang et al. 2016), and *in vitro* bound targets for 388 TFs from DAP-seq(O'Malley, Huang et al. 2016). The backend structure and tools available in *ConnecTF* are species independent and built using common software (Fix. X, Web App). The source code and detailed instructions will be provided to enable others to setbup their own instance for public or private sharing of TF-centric genomic data. A key feature of *ConnecTF* is the powerful, yet intuitive, logic-based query system (Fig. X, Query), which allows users to select TFs of interest and filter the targets based on different criteria. This includes searching for targets of all, or specific TFs of interest, and filtering based on the type of experiment for the TF-target interaction. Queries can be built using the graphical interface or by typing in more complex queries. Currently, we have several analysis and visualization tools integrated into *ConnecTF* (Fig. X, Analysis and Visualization). A table of the TF-target interactions can be accessed in the browser or downloaded for offline use. Other analyses include tests for enrichment of validated cis-motifs in the targets of each TF, the overlap between the targets of each TF and the user-submitted gene list, and the overlap between the targets of each TF when compared pairwise. Networks can be visualized using the Cytoscape plugin or downloaded. Finally, if the user uploads a predicted GRN network, *ConnecTF* uses to TF-target validation data to generate area under the Precision-Recall (AUPR) curves with an interactive sliding window feature that can be used to select cutoffs or prune the predicted network(Marbach, Costello et al. 2012). T We will use *ConnecTF* in each Aim of the proposal to: (***i***) integrate the TF-regulation and TF-binding data for 150 TFs and to identify candidate *Hit-and-Run* TFs (Aim 1A), (***ii***) classify stable vs. transient TF-target interactions for targets in (Aim 2A) to be use as priors for time-series forecasting (Aim 3B), and (***iii***) identify the indirect targets of a TF1 that respond only *in planta* (Aim 2C), and chart a network path back to the TF1 using direct targets identified in cells (Aim 3A), and (***iv***) perform AUPR on time-series GRNs we produce.

Brooks, M. D., J. Cirrone, A. V. Pasquino, J. M. Alvarez, J. Swift, S. Mittal, C.-L. Juang, K. Varala, R. A. Gutierrez, G. Krouk, D. Shasha and G. M. Coruzzi (2019). "Network Walking charts transcriptional pathways for dynamic nitrogen signaling using validated and predicted genome-wide interactions." Nature Communications.

Doidy, J., Y. Li, B. Neymotin, M. B. Edwards, K. Varala, D. Gresham and G. M. Coruzzi (2016). ""Hit-and-Run" transcription: de novo transcription initiated by a transient bZIP1 "hit" persists after the "run"." BMC Genomics **17**: 92.

Fairhead, M. and M. Howarth (2015). "Site-specific biotinylation of purified proteins using BirA." Methods Mol Biol **1266**: 171-184.

Green, N. M. (1975). "Avidin." Adv Protein Chem **29**: 85-133.

Kim, J., A. B. Cantor, S. H. Orkin and J. Wang (2009). "Use of in vivo biotinylation to study protein–protein and protein–DNA interactions in mouse embryonic stem cells." Nature Protocols **4**: 506.

Kulyyassov, A., M. Shoaib, A. Pichugin, P. Kannouche, E. Ramanculov, M. Lipinski and V. Ogryzko (2011). "PUB-MS: A Mass Spectrometry-based Method to Monitor Protein–Protein Proximity in vivo." Journal of Proteome Research **10**(10): 4416-4427.

Marbach, D., J. C. Costello, R. Küffner, N. M. Vega, R. J. Prill, D. M. Camacho, K. R. Allison, A. Aderhold, R. Bonneau and Y. Chen (2012). "Wisdom of crowds for robust gene network inference." Nature methods **9**(8): 796.

Matsuda, K., T. Mikami, S. Oki, H. Iida, M. Andrabi, J. M. Boss, K. Yamaguchi, S. Shigenobu and H. Kondoh (2017). "ChIP-seq analysis of genomic binding regions of five major transcription factors highlights a central role for ZIC2 in the mouse epiblast stem cell gene regulatory network." Development **144**(11): 1948-1958.

O'Malley, R. C., S. C. Huang, L. Song, M. G. Lewsey, A. Bartlett, J. R. Nery, M. Galli, A. Gallavotti and J. R. Ecker (2016). "Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape." Cell **166**(6): 1598.

Palaniswamy, S. K., S. James, H. Sun, R. S. Lamb, R. V. Davuluri and E. Grotewold (2006). "AGRIS and AtRegNet. A Platform to Link cis-Regulatory Elements and Transcription Factors into Regulatory Networks." Plant Physiology **140**(3): 818-829.

Para, A., Y. Li and G. M. Coruzzi (2018). "muChIP-Seq for Genome-Wide Mapping of In Vivo TF-DNA Interactions in Arabidopsis Root Protoplasts." Methods Mol Biol **1761**: 249-261.

Para, A., Y. Li, A. Marshall-Colon, K. Varala, N. J. Francoeur, T. M. Moran, M. B. Edwards, C. Hackley, B. O. Bargmann, K. D. Birnbaum, W. R. McCombie, G. Krouk and G. M. Coruzzi (2014). "Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in Arabidopsis." Proc Natl Acad Sci U S A **111**(28): 10371-10376.

Pruneda-Paz, J. L., G. Breton, D. H. Nagel, S. E. Kang, K. Bonaldi, C. J. Doherty, S. Ravelo, M. Galli, J. R. Ecker and S. A. Kay (2014). "A genome-scale resource for the functional characterization of Arabidopsis transcription factors." Cell Rep **8**(2): 622-632.

Song, L., S.-s. C. Huang, A. Wise, R. Castanon, J. R. Nery, H. Chen, M. Watanabe, J. Thomas, Z. Bar-Joseph and J. R. Ecker (2016). "A transcription factor hierarchy defines an environmental stress response network." Science **354**(6312): aag1550.

Sura, W., M. Kabza, W. M. Karlowski, T. Bieluszewski, M. Kus-Slowinska, L. Paweloszek, J. Sadowski and P. A. Ziolkowski (2017). "Dual Role of the Histone Variant H2A.Z in Transcriptional Regulation of Stress-Response Genes." Plant Cell **29**(4): 791-807.

Varala, K., A. Marshall-Colón, J. Cirrone, M. D. Brooks, A. V. Pasquino, S. Léran, S. Mittal, T. M. Rock, M. B. Edwards, G. J. Kim, S. Ruffel, W. R. McCombie, D. Shasha and G. M. Coruzzi (2018). "Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants." Proceedings of the National Academy of Sciences **115**(25): 6494-6499.

Weirauch, M. T., A. Yang, M. Albu, A. G. Cote, A. Montenegro-Montero, P. Drewe, H. S. Najafabadi, S. A. Lambert, I. Mann, K. Cook, H. Zheng, A. Goity, H. van Bakel, J. C. Lozano, M. Galli, M. G. Lewsey, E. Huang, T. Mukherjee, X. Chen, J. S. Reece-Hoyes, S. Govindarajan, G. Shaulsky, A. J. M. Walhout, F. Y. Bouget, G. Ratsch, L. F. Larrondo, J. R. Ecker and T. R. Hughes (2014). "Determination and inference of eukaryotic transcription factor sequence specificity." Cell **158**(6): 1431-1443.

Yang, D.-C., J. Jin, L. Kong, Y.-Q. Meng, G. Gao, J. Luo and F. Tian (2016). "PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants." Nucleic Acids Research **45**(D1): D1040-D1045.

Chromatin Affinity Purification (ChAP) from Arabidopsis thaliana Rosette Leaves Using in vivo Biotinylation System.

Sura et al 2018.

Use of In Vivo Biotinylation for Chromatin Immunoprecipitation

Arman Kulyyassov [Muhammad Shoaib](https://currentprotocols.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Shoaib%2C+Muhammad) [Vasily Ogryzko](https://currentprotocols.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Ogryzko%2C+Vasily) 2011