## **Contents**

# Supplementary Figures



# <span id="page-0-0"></span>Supplementary Tables



### <span id="page-1-0"></span>1 Quantitative Analysis

*Model identification*. The model mirrors the construction of the ARC<sup>Sc</sup> (Supplementary Figure [6\)](#page-13-0): Just as an ARF and a promoter were added to the previously engineered AFB*|*IAA system, an additional equation for GFP (represented by *g*) was added to the previously identified mathematical model of the AFB*|*IAA system [\[1\]](#page-7-0) to obtain the following.

$$
\dot{x} = k_1 u - k_2 x
$$
\n
$$
\dot{y} = k_3 - k_4 y - k_5 xy
$$
\n
$$
\dot{g} = -k_6 g + \frac{k_7}{1 + k_8 y}.
$$
\n(1)

The time-dependent variables  $u, x, y$  and  $g$  represent the concentration of auxin input, a lumped internal state (combining the intermediate reactions involving the binding of auxin to the AFB receptor and related molecular machinery), the IAA levels, and the GFP reporter, respectively. The parameters describe the synthesis of the internal state (*k*1), the degradation and dilution of the internal state (*k*2), expression of IAA (*k*3), degradation and dilution of IAA  $(k_4)$ , auxin-induced degradation of IAA  $(k_5)$ , degradation and dilution of GFP  $(k_6)$ , non-repressed expression of the reporter  $(k_7)$ , and the repression on the reporter  $(k_8)$ . The first two equations in the model are identical to the model of the AFB*|*IAA system reported previously [\[1\]](#page-7-0). The third equation captures the dynamics of GFP output – the basal degradation and dilution rate, and the IAA dependent expression rate. Because IAAs bind with ARF and inhibit the activation of GFP expression, the GFP expression rate is inversely proportional to the amount of IAA. The model is not intended to be mechanistic. Nevertheless, even without additional terms modeling ARF dynamics, for example by directly modeling the inhibitory effect of IAA on GFP expression, the model fits the wide range of GFP induction responses (Figure 2). In fact, the addition of a species representing the concentration of ARF introduces uncertainty to the parameter estimation that cannot be resolved because of the low number of observable outputs.

*Parameter-Component dependency*. The main question we answer in our analysis of the model is: which of the parameters  $k_1$  through  $k_8$  are tuned by the choice of AFB or IAA. We call the proteins that make up the circuit,

and in particular the ARF and IAA, *components*. As can be seen from the data, the auxin response clearly depends on the choice of components. For those aspects of the dynamics that are independent of the changing components, we expect that the model parameters associated with the components remain constant across ARC*Sc* variants. For example,  $k_1$  is a parameter associated with the choice of AFB [\[1\]](#page-7-0) and when estimating the parameters of ARC<sup>Sc</sup> circuits, all with the same AFB variant, the value for  $k_1$  is constant for all members of the set. However,  $k_3$  is associated with the choice of IAA; therefore, when estimating the parameters of ARC*Sc* circuits with different IAAs, we expect the parameter values to be different for each variant. This dependency implies that every dependent parameter becomes a variable unique to an ARC*Sc* variant (made up of unique constituent components), and the set of these unique variables add to the dimension in which we compare ARC*Sc* variants with one another. Since we seek a small set of quantitative features for simpler comparison across ARC*Sc* variants for parsimonious representation, we aim to keep the number of dependencies small.

We use a matrix to represent the relationship between parameters  $(k_1$  through  $k_8$ ) and components (the ARFs and IAAs). In particular, we define the  $m \times n$  Boolean matrix A for *m*-parameters and *n*-components. Note that *n* refers to the number of component types that compose the system, and not the number of interchangeable variants for a given component type. Here  $m = 8$  and  $n = 2$ . The entry  $a_{i,j}$  is 1 if the *i*-th parameter is dependent on the choice of the *j*-th component, and is 0, otherwise.

To validate our interpretation of the parameters  $k_1$  through  $k_8$ , we estimate  $A$  using pooled experimental data. Note that there exist  $2^{m \times n}$  candidates for *A* and each candidate is a hypothesis regarding the parameter-component dependency. Some hypotheses perform better than others with respect to how well they fits the range of divergent system behaviors across a set of system variants. These candidates tend to have more non-zero entries because each non-zero entry effectively allows another degree-of-freedom for fitting experimental data. At the same time, the candidates with more non-zero entries ultimately result in more unique variables per system variant. Therefore, we search for an *A* that results in a low model-fit residual with small number of non-zero entries.

We implemented the model fitting and parameter-component matrix search in *Mathematica*. The code and supporting data are available for download at [http://klavinslab.org/data.html.](http://klavinslab.org/data.html) Given that there are 2<sup>16</sup> possible

candidate matrices for the 2-component ARC*Sc* circuits and the 8-parameter model, directly examining all of the candidates is prohibitively expensive in computational cost (approximately 2 hours per matrix using the *NMinimize* function in *Mathematica*). Instead, we selectively choose and evaluate candidate matrices using a greedy search algorithm (Supplementary Figure [7\)](#page-14-0). The algorithm begins with 0-matrix (*parent* matrix) and searches a set of candidate matrices that are generated by adding an additional 1 to each of the available positions (previously where it was 0) (a set of *child* matrices). Among the set, the matrix with the lowest model-fit residual is chosen. The process is iterated by updating the parent matrix with the best performer and generating a next set of matrices. Though efficient, this approach sometimes fails to consider other candidate matrices that are nearly equivalent in model-fit residual to the lowest value. Repeated searches from different initial conditions, however, gave the same results in this case. Additionally, we observed a trend where the model-fit residual monotonically decreases with increasing number of non-zero entries. We conjecture that this behavior will be observed in general, but no analytical proof yet exists. Finally, it is notable that because of various uncertainties in measurement/experimental noise and estimation errors, the residual eventually saturates at some low value and further addition of non-zero entries do not decrease it further. We defined an arbitrary threshold at which point the iterative search for optimal *A* halts.

The optimal matrix *A*<sup>★</sup> (Supplementary Figure [7\)](#page-14-0) suggests the following interpretations regarding the ARC<sup>Sc</sup> and the model parameter.

- $-$  The expression strength of IAA  $(k_3)$  is one of the strongest determinants of the ARC<sup>Sc</sup> response dynamics as it was consistently chosen as the first non-zero entry in multiple iterations of greedy search algorithm. Additionally, the parameter is dependent on the choice of IAA alone, and is independent of the ARF.
- $-$  The basal expression of GFP  $(k_7)$  is dependent on the choice of ARF, but independent of the choice of IAA.
- $\sim$  Somewhat unexpectedly, the auxin-mediated IAA degradation  $(k_5)$  is dependent on the choice of both IAA and ARF. One hypothesis is that ARF and AFB compete for binding with IAA and when IAA is bound to ARF, AFB cannot bind – inhibiting auxin-mediated IAA degradation.
- $-$  The affinity of IAA to ARF  $(k_8)$  is dependent on the choices of IAA and ARF, which is consistent with the

model interpretation.

The findings of parameter-component dependency identified here could have been obtained, to some degree, through qualitative observation of  $ARC^{Sc}$  responses (e.g. The minimal dependency of  $k_3$  on the choice of ARF could have been elicited from noticing that the preauxin steady state of ARC<sup>Sc</sup> remains constant when a different ARF is used). However, by using the approach discussed here, we investigated different hypotheses and determined that not all of the model parameters need be varied to account for the range of dynamic responses in ARC*Sc*. Using the approach, we have identified a minimal set of quantities that identify the biological functions of ARC*Sc*. Furthermore, these values may be tuned by choosing the appropriate ARF and IAA.

*Sensitivity Analysis*. We defined two performance metrics to capture the overall behavior of the system: The preauxin steady state and the activation time. These metrics enable quantitative comparisons among the ARC*Sc* variants (and may guide later rational engineering of synthetic gene networks using the synthetic system analyzed here).

The preauxin steady state. The analytical expression for the preauxin steady state with  $u = 0$  is

$$
g_0 = \frac{k_7 k_4}{k_6 (k_4 + k_3 k_8)}.
$$
 (2)

Experimentally, it was shown that the maximal GFP intensity is dependent on the choice of ARF. Therefore, to fairly compare preauxin steady states of GFP among the ARC*Sc* variants with different ARFs, the values were normalized by  $k_7/k_6$ , which corresponds to the maximal GFP expression rate when *y* is zero. Thus the normalized value represents the relative fraction of the preauxin steady states of GFP with respect to the maximal GFP intensity possible for the given ARF. Note that the relative rankings of IAAs in ARF19 and ARF7 are conserved for this metric (Supplementary Figure [8\)](#page-15-0).

Sensitivity analysis reveals how the preauxin steady state changes with a change in the model parameters (Figure  $3C$  – specifically  $k_3$ ,  $k_5$ , and  $k_8$ . Each parameter was varied one at a time from the minimum and the maximum estimated values. The preauxin steady state depends on the synthesis rate of IAA and the affinity of IAA to ARF. At the same time, because the preauxin steady state is measured before auxin is added to the system, the quantity

is independent of the auxin-induced IAA degradation  $(k<sub>5</sub>)$ . Therefore, the preauxin steady state does not vary as  $k_5$  is changed, and the sensitivity analysis consistently predicts this behavior. The effect of varying  $k_3$  and  $k_8$  is qualitatively equivalent – when either of the value is decreased, the preauxin steady state increases, and vice versa.

The activation time,  $\Delta Tt$ . We define the *activation time* as the difference in time between when 10% GFP expression intensity  $(t_{10})$  and 90% GFP expression  $(t_{90})$  are reached. We set the initial state (refers to the preauxin steady state GFP intensity) as the 0% and the final state (refers to the post-auxin steady state GFP intensity) as 100%. The ODEs are solved to numerically approximate the activation time, because there is no analytical solution to the model. As in the sensitivity analysis of the preauxin steady state, each parameter was varied between the minimum and the maximum of the estimated parameter values from the data, while the other parameters are held constant (Figure 3D). The most striking trend is that the activation time is predicted with high accuracy by  $k_5$  alone and that *k*<sup>3</sup> or *k*<sup>8</sup> are poor predictors of the activation time. This suggests that when engineering a synthetic gene network using the ARC*Sc* circuit which requires faster/slower activation rate, it is far more advantageous to select an IAA with a higher/lower  $k_5$  than one with a different  $k_3$  or  $k_8$ .

The sensitivity analysis also suggests that varying  $k_3$  and  $k_8$  have similar effect on the  $\Delta T$ . This raises the possibility that these parameters are indistinguishable when the output measurement is limited to the variable *g*. However, the ambiguity is resolved if the output measurements are extended to the variable *y*, because it allows us to estimate  $k_3$  – the expression rate of IAA – independently of the  $k_8$ . To achieve a similar effect, we pooled two data sets - GFP induction data and YFP-IAA degradation data. Though the YFP-IAA degradation data is collected from a variation of ARC*Sc* circuit that lacks the plant promoter, it is a close approximation of the standard ARC*Sc* circuit. This condition allowed us to deduce the approximate range of  $k_3$  for ARC<sup>Sc</sup> circuits. Though mathematically indistinguishable in sensitivity, different IAAs with different  $k_3$  and  $k_8$  values have varying effects on the circuit performance. In particular, the scatter plot of  $k_3$  and  $k_8$  shows little correlation, suggesting that the two parameters are controlled independently within the IAA sequence and have partially independent roles in output dynamics (Supplementary Figure [9\)](#page-16-0). We hypothesize that  $k_3$  is determined by the transcriptional and translational efficiency of the IAA in question, whereas  $k_8$  is determined by the affinity of the IAA for its target ARF.

*Competition of multiple IAAs*. Instead of modifying the model structure and introducing new variables and parameters to represent the secondary IAA, we reinterpret the standard model<sup>1</sup>. Specifically, the variable  $y$  was interpreted as the *effective* IAA. This is because without engineering the multiple-IAA ARC*Sc* to measure either or both of the IAA dynamics, it is difficult to decouple the effects of individual IAAs. Therefore, we assume that there exists a lumped species whose behavior is a combination of behaviors of the two IAAs.

As discussed, the characteristic features of an IAA in ARC<sup>Sc</sup> are quantified by  $k_3$ ,  $k_5$  and  $k_8$ . Similarly, we assume the characteristic features of the effective IAA species of a mixed-IAA ARC*Sc* are quantified by the same parameters. We further hypothesize that  $k_5$  and  $k_8$  of the mixed IAA circuit are linear combinations of the parametric values of each individual IAA with the weighting coefficient parametrized by  $\alpha$  and  $\beta$ , respectively. These parameters approximate the relative competitiveness in AFB- and ARF-binding of one of the two IAA in a mixed ARC*Sc* system. This relationship is written as follows.

$$
k_{5,\text{mixed-IAA}} = \alpha k_{5,\text{IAA}_x} + (1-\alpha) k_{5,\text{IAA}_y} \tag{3}
$$

$$
k_{8,\text{mixed-IAA}} = \beta k_{8,\text{IAA}_x} + (1-\beta) k_{8,\text{IAA}_y}.
$$
 (4)

It was shown that mixed-IAA response dynamics often mimic the response of one of the individual IAA ARC*Sc* circuit more closer than the other. To determine whether this trend is bolstered by either a strong relative competitiveness of AFB- or ARF-binding of a given IAA, we simulated the effect of varying  $\alpha$  and  $\beta$  simultaneously. To examine the preferential response dynamics of a mixed ARC<sup>Sc</sup> system, we define a metric  $\omega$ , that quantifies how closely the mixed  $ARC^{Sc}$  mimics the  $IAA<sub>r</sub>$  behavior, as follows.

$$
\omega = \frac{\int |g(\theta_{xy}) - g(\theta_y)| dt}{\int |g(\theta_x) - g(\theta_y)| dt}, \qquad (5)
$$

where  $g(\theta_{xy})$  is the response dynamics of a mixed ARC<sup>Sc</sup> that co-expresses IAA<sub>x</sub> and IAA<sub>y</sub>. IAA<sub>x</sub> has near 0 dominance, when  $\alpha$  and  $\beta$  are both zeros, and complete dominance when they are both ones. With a low  $\beta$  (e.g.  $\beta$  =

<span id="page-6-0"></span><sup>&</sup>lt;sup>1</sup>The estimated parameters for the IAA-competition  $ARC<sup>Sc</sup>$  variants are shown in the Supplementary Table [4.](#page-0-0)

0, IAA<sub>x</sub> entirely loses the competition over ARF-binding to IAA<sub>y</sub>), IAA<sub>x</sub> can still dominate the response by having a high relative competitiveness in AFB-binding  $(\alpha)$ . With a low  $\alpha$ , IAA<sub>x</sub> cannot dominate even with maximum  $\beta$  value. It is not surprising that competitiveness is better captured by  $\alpha$  since  $\alpha$  corresponds to  $k_5$  which affects the auxin response much more than  $k_8$ . Note that one cannot conclude from this analysis that a faster or slower  $k_5$ predicts competition. Rather, we simply use  $\alpha$  to describe which of two IAAs dominate by showing which effective *k*<sup>5</sup> results from the competition.

### <span id="page-7-0"></span>**References**

[1] K. Havens *et al.*, A synthetic approach reveals extensive tunability of auxin signaling. *Plant Physiology*. 160(1) 135-142 (2012).

<span id="page-8-0"></span>

Supplementary Figure 1: TPL co-expression with IAA14 does not inhibit ARF activation. Full length (FL) TPL was either co-expressed with IAA14 (dark grey) or fused to IAA14 (light blue). Both constructs exhibited similar levels of preauxin steady state GFP reporter activity as ARC strains with IAA14 alone (NO RD). The other RD constructs are shown for reference. Error bars represent standard deviation (n=2).

<span id="page-9-0"></span>

Supplementary Figure 2: Flow cytometry data for all ARC*Sc* strains tested. In addition, data is shown for IAA17.t2 (grey), which lacks the ARF interaction domain, IAA1.1 (light blue square), which is a yeast codon-optimized version of IAA1 and slr (blue square), which is an auxin-insensitive form of IAA14. The top panels show mock and auxin treated data for each IAA in the presence of ARF19, while the bottom panels show the same IAAs in the presence of ARF7.

<span id="page-10-0"></span>

Supplementary Figure 3: Preauxin steady state GFP reporter activity of IAA 1.1 is lower than that of IAA1 in ARF19 circuits. IAA1.1 is codon optimized for yeast. It has a higher expression level than IAA1, while maintaining a similar rate of auxin-induced degradation [\[1\]](#page-7-0). Error bars represent standard deviation (n=2).

<span id="page-11-0"></span>

Supplementary Figure 4: IAA expression levels do not correlate with dominance behavior in ARC*Sc* competition strains. qPCR was used to measure expression of each IAA present in replicate 12+3 or 14+28 strains. In both circuits, there were only modest differences in expression between the IAAs found in each strain. Notably, the dominant IAA in each of these competition circuits is not the IAA expressed at the highest level. Error bars represent standard error (n=3).

<span id="page-12-0"></span>

Supplementary Figure 5: Assembly of pGP4GY-ccdB. The construction of pG4GY-ccdB illustrates the minimum modifications needed to generate a single integrating derivative of a pAG vector. pG4GY-ccdB was generated in a single Gibson reaction using four PCR products from a pAG template (pAG304GPD-EYFP-ccdB) and one from a yeast chromosomal template. PmeI sites were inserted via PCR primer overhangs. The Gateway<sup>®</sup> expression cassette region (black) contains the ccdB counter-selectable marker and chloramphenicol resistance cassette flanked by attR recombination sites to facilitate Gateway<sup>®</sup> LR reactions. The second targeting domain (green) begins at the first bp on the chromosome following the sequence homologous to the TRP1 selectable marker cassette (blue); the TRP1 coding region starts at bp +1. This figure is not to scale.

<span id="page-13-0"></span>

Supplementary Figure 6: The block diagram of the ARC*Sc* and its model. (A) The subcomponent, AFB*|*IAA, was characterized previously. The ARF and GFP components were added here. The output of the first component (IAA), interacts with the ARF and affects the system output (GFP). (B) The mathematical model is extended to encompass the dynamics of the ARC*Sc* by adding a third equation for the GFP dynamics. The block diagram illustrates the analogous composition of both the physical system and the mathematical model identification presented.

<span id="page-14-0"></span>

Supplementary Figure 7: Model-fit residual decreases with fewer constraints on parameter estimation. The increasing number of non-zero entries in the Parameter-Component dependency matrix corresponds to fewer constraints on the optimization of parameters to fit experimental data. The figure shows an instance of the greedy search algorithm, where monotonically decreasing behavior is observed. We conjecture that this behavior will be observed in general, but no analytical proof yet exists.

<span id="page-15-0"></span>

Supplementary Figure 8: Preauxin steady state values of ARC variants. A total of 20 different ARCs – 2 different ARFs in combination with 10 different IAAs – were evaluated. They are ordered according to GFP intensity in the ARF19 ARCs. For each ARC*Sc*, two to four replicates were analyzed and the error bar shows one standard deviation among the replicates.

<span id="page-16-0"></span>

Supplementary Figure 9:  $k_3$  and  $k_8$  are not correlated. Scatter plot of  $k_3$  and  $k_8$  values are plotted for the 11 IAAs assayed with ARF19. Two to four biological replicates were analyzed per ARC, and each point represents the mean of the estimated parameter across replicates. The dashed lines indicate the mean of  $k_3$  and  $k_8$  across all IAAs.

#### **Supplementary Table 1**: Yeast strains used in this study





#### **Supplementary Table 2**: Primers used in this study

#### **Primers to clone ARFs into pDONR**



#### **Adapter primers used to increase the att sites on cloned ARF PCR product**



#### **Primers for Gibson assembly of pGP4G-TPL-N100-ccdB**



PCRs fused in Gibson prNL1386 with prNL1376 using A.t. cDNA as template prNL1481 with prKL559 using pGP4G-ccdB as template prKL1390 with prKL771 using pGP4G-ccdB as template

#### **Primers for Gibson assembly of pGP4G-TPL-N300-ccdB**



PCRs fused in Gibson

prNL1386 with prNL1377 using A.t. cDNA as template prNL1481 with prKL559 using pGP4G-ccdB as template prKL1390 with prKL771 using pGP4G-ccdB as template







### **Supplementary Table 4.** Estimated parameters of competition Auxin Response Circuits