SUPPLEMENTARY MATERIAL
for
Conserved and differential gene interactions in dynamical biological systems

Zhengyu Ouyang, Mingzhou Song, Robert Gueth, Thomas J. Ha,
Matt Larouche and Dan Goldowitz

June 27, 2011

Contents

1 Limitations of differential correlation and numerical comparison

We demonstrate the limitations of differential correlation and numerical comparison in a linear model. Let two random variables, X and Y, have a linear relationship $Y = kX + c + \epsilon$, where slope k and intercept c are constants, and ϵ a normal noise $N(0, \sigma^2)$. We first establish the mathematical relationship between model coefficients and the correlation coefficient. Give n observations of X and Y, $(x_1, y_1), ..., (x_n, y_n)$, the Pearson correlation coefficient between X and Y is defined as

$$r = \frac{1}{n}\sum_{i=1}^{n}[\frac{x_i - \bar{x}}{\sigma_x}\frac{y_i - \bar{y}}{\sigma_y}] \tag{1}$$

where $\bar{x}$ and $\bar{y}$ are the sample means, σ_x and σ_y are the sample standard deviations, of X and Y, respectively. Given the linear model, we derive

$$\bar{y} = k\bar{x} + c \tag{2}$$

and

$$\sigma_y = \sqrt{k^2\sigma_x^2 + \sigma^2} \tag{3}$$

Then we have

$$r = \frac{k\sigma_x}{\sqrt{k^2\sigma_x^2 + \sigma^2}} \tag{4}$$

Now we consider comparison of two unknown models $Y = k_1X + c_1 + \epsilon_1$ and $Y = k_2X + c_2 + \epsilon_2$, where ϵ_1 and ϵ_2 are two noise terms following $N(0, \sigma_1^2)$ and $N(0, \sigma_2^2)$, respectively. Let r_1 and r_2 be correlation coefficients of data collected from each model, respectively. The differential correlation score dc is defined by

$$dc = r_1 - r_2 = \frac{k_1}{\sqrt{k_1^2 + \sigma_1^2/\sigma_{x_1}^2}} - \frac{k_2}{\sqrt{k_2^2 + \sigma_2^2/\sigma_{x_2}^2}} \tag{5}$$

In the differential correlation strategy for comparing interactions, one test the hypothesis $k_1 \neq k_2$ for model difference by inspecting $dc \neq 0$. However, using dc is not always statistically powerful. Without any additional assumptions, $dc \neq 0$ is neither a necessary nor a sufficient condition for $k_1 \neq k_2$. When the noise variances σ_1^2 and σ_2^2 are negligible relatively to the data variance, dc will approach 0, -2, or 2. In this case, differential correlation can indeed capture the sign difference, but fails to capture any numerical difference between k_1 and k_2. In more general situations, there are two issues with differential correlation. The first issue is that data variances $\sigma_{x_1}^2$ and $\sigma_{x_2}^2$ may be different across conditions, which can happen when the systems under comparison have different strength of dynamic variation (not due to noise) across conditions, likely to happen in biological experiments. The second issue is that the noise variances may also be different across conditions, when it is difficult to maintain the same noise level in experimentation. These issues are often addressed in part by proper normalization, but can be a challenge when comparing models using data generated from different sources.

The above two issues are ignored by assumptions in testing $k_1 \neq k_2$ by $dc \neq 0$. That is, the data variances must be assumed to be equal across conditions $\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_x^2$, and so do the noise variances $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Under these assumptions, we have

$$dc = \frac{k_1}{\sqrt{k_1^2 + \sigma^2/\sigma_x^2}} - \frac{k_2}{\sqrt{k_2^2 + \sigma^2/\sigma_x^2}} \tag{6}$$

Only under these restrictive assumptions, $dc \neq 0$ is both a necessary and sufficient condition for $k_1 \neq k_2$.

Under one more assumption of equality between data and noise variances, $\sigma_x^2 = \sigma^2$, differential correlation reduces to numerical comparison or reconstruct-then-compare. This is also a highly restrictive assumption. Mathematically it leads to

$$dc = \frac{k_1}{\sqrt{k_1^2 + 1}} - \frac{k_2}{\sqrt{k_2^2 + 1}} \tag{7}$$

It is unlikely that a real-world experiment can satisfy all these restrictive assumptions. Therefore, differential correlation and numerical comparison can only partially address the model comparison problem, under strong and often-unpractical assumptions.

2 The heterogeneity F-test generalizes differential correlation and numerical comparison

The F-test used in regression analysis can assess how a linear model $Y = kX + c$ explains the relationship between two random variables X and Y better than a zero slope null model $Y = c$. Let $(x_1, y_1), ..., (x_n, y_n)$ be samples of size n from (X, Y). Let RSS_d be the residual sum of squares between observations, y_i, and model predictions, $\hat{y}_i$, made by an optimal estimator. Let RSS_n be the total residual sum of squares over y_i, relative to the null model. Regardless of model linearity, the F-statistic is defined by

$$F = \frac{RSS_n - RSS_d}{RSS_d/(n-2)} \tag{8}$$

where F asymptotically follows an F-distribution with 1-numerator and $(n-2)$-denominator degrees of freedom under the null hypothesis of $Y = c$.

When the linear model $Y = kX + c + \epsilon$, with ϵ being a random normal noise $N(0, \sigma^2)$, is used, we have

$$RSS_n = \sum_{i=1}^{n} (y_i - \bar{y})^2 = n\sigma_y^2 = n(k^2\sigma_x^2 + \sigma^2) \tag{9}$$

and

$$RSS_d = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \epsilon_i^2 = n\sigma^2 \tag{10}$$

Plugging RSS_n and RSS_d into Eq. (8), we obtain

$$F = (n-2)\frac{k^2\sigma_x^2}{\sigma^2} = \frac{n-2}{\frac{1}{r^2} - 1}$$

where r is the sample correlation coefficient between X and Y. This equation suggests that F is equivalent to r given the sample size, but F is insensitive to the sign of r or the polarity of the correlation.

Now we consider comparison of two unknown models, $Y = k_1X + c_1 + \epsilon_1$ and $Y = k_2X + c_2 + \epsilon_2$, using two observed data sets from each model. Noise terms ϵ_1 and ϵ_2 are two random variable following $N(0, \sigma_1^2)$ and $N(0, \sigma_1^2)$, respectively. One could compute F_1 and F_2 statistics for each data set and then determine if they are equivalent by testing $F_1 - F_2 = 0$. Evidently, this would become even more restrictive than dc which compares r_1 and r_2, not r_1^2 and r_2^2.

Instead, we adopt a different strategy to compare two unknown models using the heterogeneity F test. To test any difference between the models $k_1 \neq k_2$ or $c_1 \neq c_2$, the idea is to examine whether using two different models, one for each data set, achieves better fitting, than using a shared model for both data sets. This is known as the heterogeneity test (Zar, 2009). Let $(x_i^{\binom{1}{2}}, y_i^{\binom{1}{2}})$ indicate data point i in the pooled data set. We call the two separate models the heterogeneous model defined by

$$y_i^{\binom{1}{2}} = \begin{cases} y_i^{(1)} = k_1 x_i^{(1)} + c_1 + \epsilon_{1i} \text{ if } (x_i^{(1)}, y_i^{(1)}) \in \text{data set 1}, \\ y_i^{(2)} = k_2 x_i^{(2)} + c_2 + \epsilon_{2i} \text{ if } (x_i^{(2)}, y_i^{(2)}) \in \text{data set 2}. \end{cases} \tag{11}$$

and we call the shared model the homogeneous model defined by

$$y_i^{\binom{1}{2}} = k x_i^{\binom{1}{2}} + c + \epsilon_i \tag{12}$$

for all data points regardless of the data set. Then we test the inequality of the two unknown linear models by the heterogeneity F-statistic

$$F_d = \frac{(RSS_{hm} - RSS_{ht})/2}{RSS_{ht}/(n_1 + n_2 - 4)} \tag{13}$$

where RSS_{hm} and RSS_{ht} are the residual sum of squares between observations and predictions by the homogeneous model and the heterogeneous model, respectively, and n_1 and n_2 are sample sizes of the two data sets. F_d asymptotically follows an F distribution with 2-numerator and $(n_1 + n_2 - 4)$-denominator degrees of freedom under the null hypothesis that the two models are identical (Zar, 2009).

We can further derive that

$$RSS_{ht} = \sum_{j=1}^{2}\sum_{i=1}^{n}(y_i^{(j)} - \hat{y}_i^{(j)he})^2 = n_1\sigma_1^2 + n_2\sigma_2^2 \tag{14}$$

$$RSS_{hm} = \sum_{j=1}^{2}\sum_{i=1}^{n}(y_i^{(j)} - \hat{y}_i^{(j)ho})^2 \tag{15}$$

$$= n_1\sigma_1^2 + n_2\sigma_2^2 + \frac{n_1n_2(k_1-k_2)^2\sigma_{x_1}^2\sigma_{x_2}^2 + \frac{n_1^2n_2}{n_1+n_2}[(k_2-k_1)\mu_{x_2} - C]^2\sigma_{x_1}^2 + \frac{n_1n_2^2}{n_1+n_2}[(k_2-k_1)\mu_{x_1} - C]^2\sigma_{x_2}^2}{n_1\sigma_{x_1}^2 + n_2\sigma_{x_2}^2 + \frac{n_1n_2(\mu_{x_1}-\mu_{x_2})^2}{n_1+n_2}} \tag{16}$$

where $\hat{y}_i^{(j)ht}$ and $\hat{y}_i^{(j)hm}$ are predictions of heterogeneous and homogeneous models for the i-th data point, respectively, and $C = c_1 - c_2$. Then we finally obtain

$$F_d = \frac{n_1 + n_2 - 4}{2(n_1\sigma_1^2 + n_2\sigma_2^2)} \frac{n_1n_2(k_1-k_2)^2\sigma_{x_1}^2\sigma_{x_2}^2 + \frac{n_1^2n_2}{n_1+n_2}[(k_2-k_1)\mu_{x_2} - C]^2\sigma_{x_1}^2 + \frac{n_1n_2^2}{n_1+n_2}[(k_2-k_1)\mu_{x_1} - C]^2\sigma_{x_2}^2}{n_1\sigma_{x_1}^2 + n_2\sigma_{x_2}^2 + \frac{n_1n_2(\mu_{x_1}-\mu_{x_2})^2}{n_1+n_2}} \tag{17}$$

where F_d considers both data and noise variances and checks the equality of k_1 and k_2 as well as c_1 and c_2. In this definition, $F_d \neq 0$ is a necessary and sufficient condition for the two models to be different, i.e., either $k_1 \neq k_2$ or $c_1 \neq c_2$.

Now we establish the connection among the heterogeneity test, differential correlation and numerical comparison. As the latter two are independent of sample size, we assume $n_1 = n_2 = n$. As differential correlation ignores the intercept, we assume equal intercepts ($C = 0$) to focus on comparison of the slopes. With these assumptions, we have

$$F_d = \frac{n-2}{\sigma_1^2 + \sigma_2^2} \frac{(k_1-k_2)^2\sigma_{x_1}^2\sigma_{x_2}^2 + \frac{1}{2}(k_1-k_2)^2\mu_{x_2}^2\sigma_{x_1}^2 + \frac{1}{2}(k_1-k_2)^2\mu_{x_1}^2\sigma_{x_2}^2}{\sigma_{x_1}^2 + \sigma_{x_2}^2 + \frac{1}{2}(\mu_{x_1} - \mu_{x_2})^2}. \tag{18}$$

And with an assumption of $\mu_{x_1} = \mu_{x_2} = \mu_x$, we get

$$F_d = \frac{(n-2)(k_1-k_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\left(\frac{2\sigma_{x_1}^2\sigma_{x_2}^2}{\sigma_{x_1}^2 + \sigma_{x_2}^2} + \mu_x^2\right) \tag{19}$$

With three more assumptions of $\mu_x = 0$, $\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_x^2$ and $\sigma_1^2 = \sigma_2^2 = \sigma^2$, we finally obtained

$$F_d = \frac{n-2}{4}\left(\frac{k_1\sigma_x}{\sigma} - \frac{k_2\sigma_x}{\sigma}\right)^2 = \frac{n-2}{4}\left(\frac{r_1}{\sqrt{1-r_1^2}} - \frac{r_2}{\sqrt{1-r_2^2}}\right)^2 \tag{20}$$

This last form suggests that F_d becomes equivalent to dc under all the above assumptions. As F_d and its p-value is one-to-one mapping, its p-value also becomes equivalent to dc under assumptions of $n_1 = n_2 = n$, $c_1 = c_2$, $\mu_{x_1} = \mu_{x_2} = 0$, $\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_x^2$ and $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Under one more restrictive assumption, $\sigma_x^2 = \sigma^2$, it follows that F_d reduces to numerical comparison. Therefore, F_d is the most general statistics for testing inequality of two unknown linear models.

3 Reconstruction of dynamical system models

A dynamical system model (DSM), consisting of a set of ordinary differential equations (ODE) and defined in the main text, can be reconstructed from sufficient observations of a system. Given observed time course data and the form of f_i, the coefficient vector $\boldsymbol{\beta}_i$ can be estimated for each variable. We formulate the estimation problem as multiple linear regression and solve it by the least squares method. Let matrix $X = (\mathbf{x}[1]^\top, \mathbf{x}[2]^\top, ..., \mathbf{x}[T]^\top)$ contains time course data collected from the system vector $\mathbf{x}$ at T time points. Let $y_i[t]$ denote the change rate of a variable x_i at time t as

$$y_i[t] = \frac{dx_i[t]}{dt}. \tag{21}$$

We estimate this derivative from time course observation by using a smoothing *spline* technique. Let a vector $\mathbf{y}_i = (y_i[1], y_i[2], ..., y_i[T])^\top$ represent the estimated derivatives of x_i at the T time points. With a given form of f_i, the coefficient vector $\boldsymbol{\beta}_i$ can be estimated by multiple linear regression. Let $\hat{\boldsymbol{\beta}}_i$ be the estimator. One can test whether $\boldsymbol{\beta}_i$ deviates from $\mathbf{0}$, by the F-statistic defined as

$$F = \frac{(|\mathbf{y}_i - \bar{y}_i|^2 - |\hat{\mathbf{y}}_i - \mathbf{y}_i|^2)/(\dim(\hat{\boldsymbol{\beta}}_i))}{|\hat{\mathbf{y}}_i - \mathbf{y}_i|^2/(T - \dim(\hat{\boldsymbol{\beta}}_i) - 1)} \tag{22}$$

where

$$\bar{y}_i = \frac{1}{T}\sum_{t=0}^{T-1} y_i[t] \tag{23}$$

is the mean of the derivative over the time course, and also the model prediction under null hypothesis with one free parameter. Under the alternative hypothesis, the predicted derivative is

$$\hat{\mathbf{y}}_i = f_i(\mathbf{X}, \hat{\boldsymbol{\beta}}_i) + \beta_{i0} \tag{24}$$

using the estimated model with ($\dim(\hat{\boldsymbol{\beta}}_i)$+1) free parameters. Under a normal noise distribution, the F-statistic asymptotically follows an F-distribution, with $\dim(\hat{\boldsymbol{\beta}}_i)$ numerator degrees of freedom and $(T - \dim(\hat{\boldsymbol{\beta}}_i) - 1)$ denominator degrees of freedom, under the null hypothesis of $\boldsymbol{\beta}_i = 0$. The p-value, the upper-tail probability of the F-statistic, indicates how significantly $\hat{\boldsymbol{\beta}}_i$ deviates from 0. Thus the p-value translates to the goodness-of-fit of f_i to the observed data.

In addition to indicating the goodness-of-fit to the observations for a given f_i, the F-statistic provides a criterion to select an appropriate f_i. A choice of the form for f_i must specify its mathematical formula and identities of involved variables, but not the coefficient values. The complexity of f_i is measured by its number of free variables, related to the number of independent variable in f_i. The goodness-of-fit is how well f_i, with an estimated coefficients $\hat{\boldsymbol{\beta}}_i$ and $\hat{\beta}_{i0}$, approximates observed x_i, typically measured by a mean square distance between prediction and observation. In the F-statistic defined in Eq. (22) and its p-value, both the goodness-of-fit and the complexity of f_i are accounted for. Thus, the F-statistic provides a criterion to test a given f_i with a specific form and its estimated coefficients.

4 Comparative DSM p-value correction by permutation tests

We use permutation tests to approximate the null distributions and adjust for multiple testing effects in the comparative DSM (CDSM). The three test statistics follow F-distributions with three assumptions: First, the sample size is sufficiently large; second, observations at different time points are independent; third, the noise is normally distributed. These assumptions are not always met in biological experiments. The sample size is often relatively small to the number of variables. Observations over sampling time points may not be independent in a dynamical system. Moreover, noise due to biological variation does not necessarily follow a normal distribution. Therefore, we use permutation tests to correct the null distribution of the three test

statistics. First, to obtain the expression rate of change, we still use *pspline* (Heckman and Ramsay, 1996) to obtain the derivatives for each conditions, $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$. Second, to eliminate true differential interactions, we pool data from different conditions $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, to form $\mathbf{x}^{(1,2)}$. This merging happens for each variable separately. Third, to eliminate true interactions, we permute the pooled time course data, $\mathbf{x}^{(1,2)}$, variable by variable to obtain $\mathbf{x}_p^{(1,2)}$. Fourth, to eliminate any true conserved interactions, we split the permuted pool data according to original data sizes under different conditions, $\mathbf{X}_{perm}^{(1)}$ and $\mathbf{X}_{perm}^{(2)}$. Finally, we compute test statistics F_c, F_d and F_t from $\mathbf{X}_{perm}^{(1)}$ and $\mathbf{X}_{perm}^{(2)}$ according to equations given in the main text. Empirical distributions for each test statistic can be estimated by repeating above five steps a large number of times depending on the required precision for p-value. When multiple parents are considered, the permutation tests also adjust for family-wise false positive rate.

5 Numerical comparison or reconstruct-then-compare

Numerical comparison, also called reconstruct-then-compare, detects differential interactions by numerically comparing estimated models after reconstructing them separately in each condition. If ODEs are used to represent two systems, the distance between two coefficient vectors is used to detect interaction shifts by comparing them with a predefined threshold. Considering different parameter scales in ODEs, we define the distance between two coefficient vectors of an interaction across two conditions by

$$Score_{NC} = \max_{i=1,\ldots,k}(|\hat{\beta}_i^{(1)} - \hat{\beta}_i^{(2)}|/(2\max(\hat{\beta}_i^{(1)}, \hat{\beta}_i^{(2)})))$$

where k is the dimension of $\hat{\boldsymbol{\beta}}$. If $Score_{NC}$ is greater than a given threshold, a differential interaction is detected. This strategy works well only when uncertainty in the estimated parameters can be ignored.

6 Differential correlation

Differential correlation is a strategy to detect a differential interaction between two variables by comparing their correlation coefficients across conditions. It is the most widely used method for differential interaction analysis so far. One method for differential gene-gene correlation detects changes in the whole correlation vector associated with each gene (Hu *et al.*, 2009). We use this method in our simulation studies and compare its performance with our CDSM framework. The differential correlation measures change in gene-gene interaction by

$$\begin{aligned} Score_{DC} =& \frac{2}{n_s^2}\sum_{l_1=1}^{n_s}\sum_{l_2=1}^{n_s}\sqrt{\sum_{j=1}^{G}(w_{ij}^{(1,l_1)} - w_{ij}^{(2,l_2)})^2} \\ &- \frac{1}{n_s^2}\sum_{l_1=1}^{n_s}\sum_{l_2=1}^{n_s}\sqrt{\sum_{j=1}^{G}(w_{ij}^{(1,l_1)} - w_{ij}^{(1,l_2)})^2} \\ &- \frac{1}{n_s^2}\sum_{l_1=1}^{n_s}\sum_{l_2=1}^{n_s}\sqrt{\sum_{j=1}^{G}(w_{ij}^{(2,l_1)} - w_{ij}^{(2,l_2)})^2} \end{aligned} \tag{25}$$

where n_s is the number of independent experiments under each condition, and G is the number of genes in the network. $w_{ij}^{(1,k)}$ measures the correlation between gene i and j for k-th experiments under condition 1 and it is defined as

$$w_{ij} = \frac{1}{2}\log\frac{1+\rho_{ij}}{1-\rho_{ij}} \tag{26}$$

where ρ_{ij} is the Pearson correlation coefficient. We adopted the multivariate correlation (Zhu *et al.*, 2007) method to estimate the correlation coefficient between gene i and gene j considering the *time* and *replicate*

$d[C2]/dt$	=	$k_6[M] - k_8[\sim P][C2] + k_9[CP]$
$d[CP]/dt$	=	$-k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP]$
$d[pM]/dt$	=	$k_3[CP][Y] - k'_4[pM] - k_4[pM]([M]/[CT])^2$
$d[M]/dt$	=	$k'_4[pM] + k_4[pM]([M]/[CT])^2 - k_5[\sim P][M]$ $-k_6[M]$
$d[Y]/dt$	=	$k_1[aa] - k_2[Y] - k_3[CP][Y]$
$d[YP]/dt$	=	$k_6[M] - k_7[YP]$

Table S1: The dynamical system model governing the *cdc-cyclin* interaction in the cell division cycle in Figure S1: t, time; k_i, rate constant; aa, amino acids. The concentrations $[aa]$ and $[\sim P]$ are assumed to be constant. Variable $[C2]$ is for *cdc2*, $[CP]$ for *cdc2-P*, $[pM]$ for *preMPF*=P-cyclin-cdc2-P, $[M]$ for active *MPF* (P-cyclin-cdc2), $[Y]$ for *cyclin* and $[YP]$ for *cyclin-P*, and $[CT]$ for total *cdc2* including isolated *cdc2* and the ones in all protein complex. The two differential interactions involved in biochemical reaction #4 are the ODEs for $[pM]$ and $[M]$, where both are influenced by the rate constant k_4 of reaction #4.

factors to compensate the Pearson correlation coefficient for its inadequacy in such experiments. A differential interaction is detected if and only if $Score_{DC}$ is greater than a predefined threshold.

7 Simulation study of a cdc2-cyclin cell division cycle model with known network architecture

In this simulation study, we compare the performance of CDSM (statistical comparison) and numerical comparison, on a biological system with *known* network architecture without values of kinetic parameters. The result was briefly summarized in a previous conference paper (Ouyang and Song, 2009), but here we give a complete description. The purpose of using known network architecture is to test the effectiveness of the F-statistics in our comparative modeling. We examine the heterogeneity of interactions for each node in a cell division cycle model across two conditions.

The cell division cycle model, involving six proteins or protein complexes, captures *cdc2-cyclin* interaction in the cell division cycle (Tyson, 1991). In the system (Figure S1), *cdc2* and *cyclin* form a protein complex called *MPF* playing a central role. Six kinetic equations (Table S1) approximate dynamical interactions among the six proteins in the system. Following recommendations for coefficient values in (Tyson, 1991), we set $k_1[aa]/[CT] = 0.015min^{-1}$, $k_2 = 0$, $k_3[CT] = 200min^{-1}$, $k'_4 = 0.018min^{-1}$, $k_5[\sim P] = 0$, $k_6 = 1min^{-1}$, $k_7 = 0.6min^{-1}$, $k_8[\sim P] = 1000000min^{-1}$ and $k_9 = 1000min^{-1}$, where $[CT]$ was assumed to be a constant of 1. Coefficient k_4 is a rate constant associated with the autocatalytic activation of *MPF* by dephosphorylation of the *cdc2* subunit (reaction #4 in Figure S1). When $k_4 < 100min^{-1}$, *MPF*, being maintained in inactive forms, enters the excitable domain (as in the resting phase of non-proliferating somatic cells). As cells grow, k_4 increases (activator accumulates) and drives the system into the oscillation domain. After cell division becomes growth controlled at $k_4 > 150min^{-1}$, *MPF* enters the oscillation domain in which it alternates between active and inactive forms with a period of 35 min, roughly the cell cycle length in early frog embryos (Tyson, 1991). The subsequent burst of *MPF* activity triggers mitosis, causes k_4 to decrease (activator degrades), and brings the system back into the excitable domain (steady-state behavior).

We set up two groundtruth DSMs for comparative modeling, by varying the rate constant k_4 of dephosphorylation of the cdc2 unit on the active MPF, as indicated in Figure S1. In the first system, k_4 was set to 180 min^{-1} representing the oscillation domain in the cell cycle model. In the second system, k_4 was randomly chosen between 70 and 80 min^{-1} representing the excitable domain in the cell cycle model. From Table S1, a change in k_4 implicates two differential interactions: one for *preMPF* (pM) and one for active *MPF* (M); all remaining interactions for the other four proteins are exactly the same or conserved. These conserved and differential interactions serve as the groundtruth for performance evaluation of our comparative modeling methods.

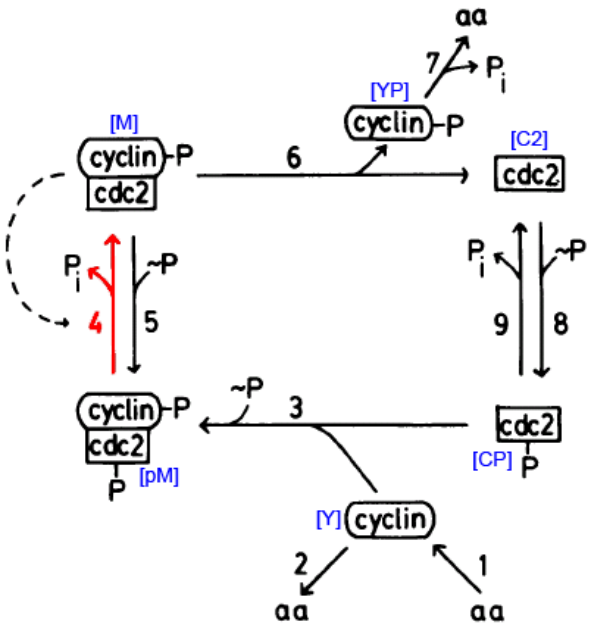

Figure S1: The *cdc2-cyclin* interaction network involved in the cell division cycle. The variable names used in the dynamical system model (Table S1) are marked next to the proteins or protein complexes they represent. Depending on the rate of the biochemical reaction marked by #4, the dynamics switches between two systems: excitable and oscillation domains. Therefore the biochemical reaction marked in red and associated with #4 contributes to differential interactions between the two systems or domains. This figure is adapted from (Tyson, 1991).

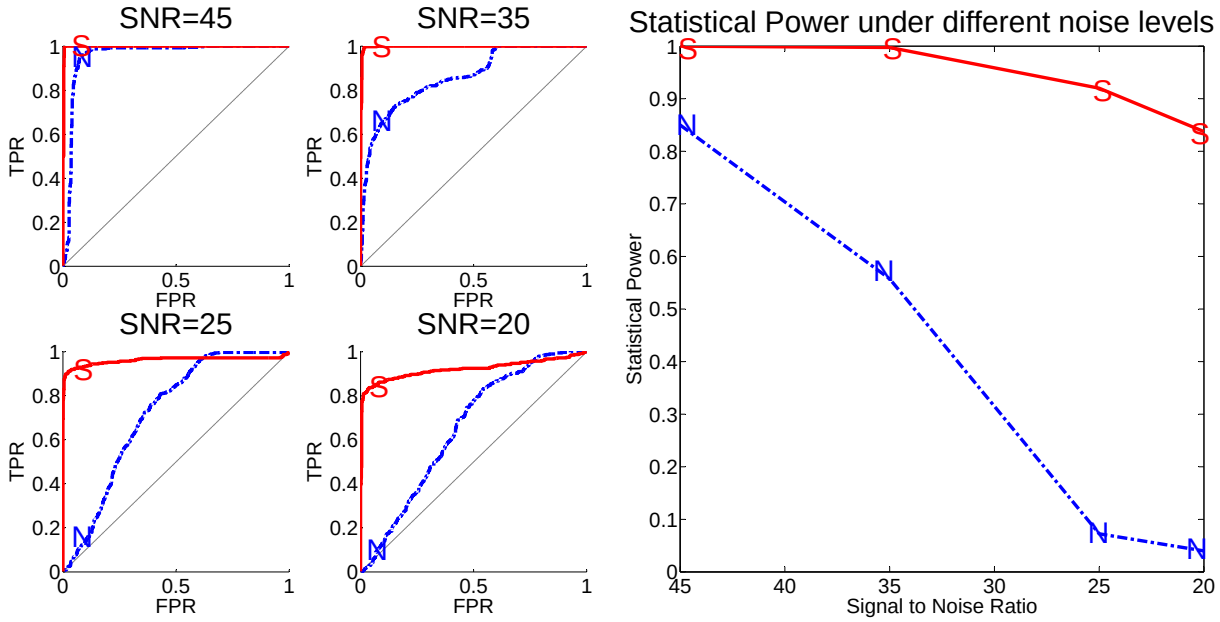

Figure S2: The advantage of CDSM modeling versus numerical comparison in detecting two differential interactions in the *cdc2-cyclin* cell division cycle model. Performance of CDSM is marked by 'S' in solid lines, and numerical comparison is marked by 'N' in dash-dotted lines. Left: the ROC curves under four noise levels. CDSM has substantial advantage over numerical comparison under the noise levels tested. Right: the statistical power of both methods as a function of the noise level. CDSM degrades much slower in performance than numerical comparison, as noise increases.

We simulated dynamic trajectories from the two groundtruth DSMs. We first simulated time courses with 20 observations over 40-min long trajectories for each domain/DSM. Then we added independent normally distributed noise three times to generate three replicates. The noise level in this study is represented by signal-to-noise ratio (SNR), defined as ten times $\log_{10}$ of the sum of squares of the signal, divided by the sum of squares of the noise. When replicates are available, the SNR can be estimated by

$$10\log_{10}\frac{\sum_{i=1}^{K}\bar{O}_{i\cdot}^2}{\sum_{i=1}^{K}(\sum_{j=1}^{M_i}(O_{ij}-\bar{O}_{i\cdot})^2)/M_i}$$

where O_{ij} is the value of replicate j of observation i, among a total of K observations, observation i contains M_i replicates, and $\bar{O}_{i\cdot}=\frac{1}{M_i}\sum_{j=1}^{M_i}O_{ij}$. We added four different levels of noise at the SNRs of 45, 35, 25, and 20.

We applied the CDSM modeling and the numerical comparison methods to detect differential interactions in the two DSMs corresponding to the two cell cycle domains. Both methods were applied to compare the interactions at each protein in the network from data generated under the four noise levels. The assumption here is that the form of and the terms in each kinetic equation are known but not the coefficients. According to $Score_{NC}$ of numerical comparison and F_d of CDSM, differential interactions can be detected. Then we compared the detection results with the two groundtruth differential interactions at $[pM]$ and $[M]$ and four ground truth conserved interactions at $[C2]$, $[CP]$, $[Y]$ and $[YP]$ to compute the overall true positive rate (TPR) and false positive rate (FPR) at each noise level.

The ROC curves generated by varying thresholds on *type I* error rate are plotted in F Figure S2 for each method. CDSM achieved consistently better performance than numerical comparison on differential interaction detection at all noise levels. When the noise was low (SNR=45dB), both methods achieved good performance. However, as the noise level increases (SNR=35, 25, 20dB), the performance of numerical

comparison drops quickly, but CDSM still maintains its power at above 80% with *type I* error rate of 0.05. As the difference between the two approaches is in the statistics used to compare interactions, the results support the effectiveness of our proposed F-statistics for comparative modeling.

8 Testing given genetic interactions for conserved and differential interactions during mouse cerebellar development

The CDSM was applied to analyze given interactions for their conserved or differential involvement during cerebellar developments. A total of 104 known genetic interactions were extracted from the BioGrid (Stark *et al.*, 2006, 2010) - a database summarizing genetic and physical interactions between genes in a variety of organisms. We are interested in knowing if these genetic interactions are conserved or differential between pre-EGL stage and EGL expansion by the proposed CDSM based on time course microarray data collected from the developing mouse cerebellum as described in the main text. Each gene interaction is represented by the following sigmoidal model:

$$\frac{dx_i(t)}{dt} = \beta_{i0} + \beta_{ij}\frac{x_j^{n_{ij}}(t)}{k_{ij}^{n_{ij}} + x_j^{n_{ij}}(t)} + \beta_{ii}\frac{x_i^{n_{ii}}(t)}{k_{ii}^{n_{ii}} + x_j^{n_{ii}}(t)} - \beta_i x_i(t) \tag{27}$$

where $x_i(t)$ is the expression level of child transcript i at time t, j is the index to the parent transcript for transcript i, β_{ij} is a constant for the influence of transcript j on transcript i, $\beta_{ii} = 0$ means no self-regulation, and $\beta_i x_i$ represents *mRNA* degradation with rate constant $\beta_i > 0$. The mathematical form of interaction used is based on (Prill *et al.*, 2010). k_{ij} is the dissociation constant between TF j and the promoter of gene i and n_{ij} the hill coefficient. We also consider self influence (the third term) with $\beta_{ii} \neq 0$. But we set $k_{ij} = k_{ii}$ and $n_{ij} = n_{ii}$ for computational simplification and automatically choose from {5,6,7,8,9,10,11} and {2,3,4,5}, respectively. Across the pre-EGL and EGL expansion, we performed comparative modeling to compute all F-statistics for each interaction, and determined if it is conserved, differential, or inactive. We applied the comparative modeling separately for each of the two mouse strains.

We identified 58 and 52 active interactions for DBA and BL6 mouse strains, respectively. All active interactions are listed in Table S2 (DBA) and Table S3 (BL6). There are 36 differential and 21 conserved interactions in the DBA strain, and 21 conserved and 30 differential interactions in the BL6 strain, at a significant p-value of less (or equal) than 0.05. Examples of differential and conserved interactions are shown by the phase diagrams in Figure S3. A strong differential interaction pattern of the pair *Meis1.1400575-Scx.130066* in the DBA strain can be observed by the diverging trajectories, while the interaction between another pair *Six3.3830402-Pax6.101660253* in the BL6 strain is consistent in transition direction and location in the two stages, demonstrating a conserved nature of the interaction. Although these detected gene interactions have been biologically determined in other organisms for various biological functions, their roles in cerebellar development are yet to be confirmed and they give rise to new hypotheses for further biological experiments. Since we have divided the microarray expression data into pre-EGL and EGL expansion time periods, which is effectively before and after granule cells are present, respectively; and granule neurons are the most abundant population of neurons in the cerebellum, our analyses might be biased towards granule cell events. Furthermore, it is believed that granule cells progenitors ramp up the rate of cell division during EGL expansion phase to produce the billions of cells that constitute this population in the mature cerebellum. Therefore it seems reasonable that genes involved in cell cycle regulation, morphogenesis and apoptosis should be regulated differentially between the two periods.

Some of the hypothetical gene interactions, computed by CDSM, have indeed been exposed by various studies of the cerebellum. The following is a summary of biological literature support for gene interactions we identified:

E2f1-E2f2. During proliferation of cerebellar granule neurons, *E2f1* and *E2f2* are activated by sonic hedgehog (Shh) signaling via *n-myc* and cyclin D1 (*cycD*) (Oliver *et al.*, 2003). Together, these three proteins - *E2f1*, *E2f2*, *cycD*) - promote proliferation when *n-myc* is over-expressed in granule cell precursors. Thus it suggests that the formation and expansion (proliferation) of the EGL could be directly linked to the

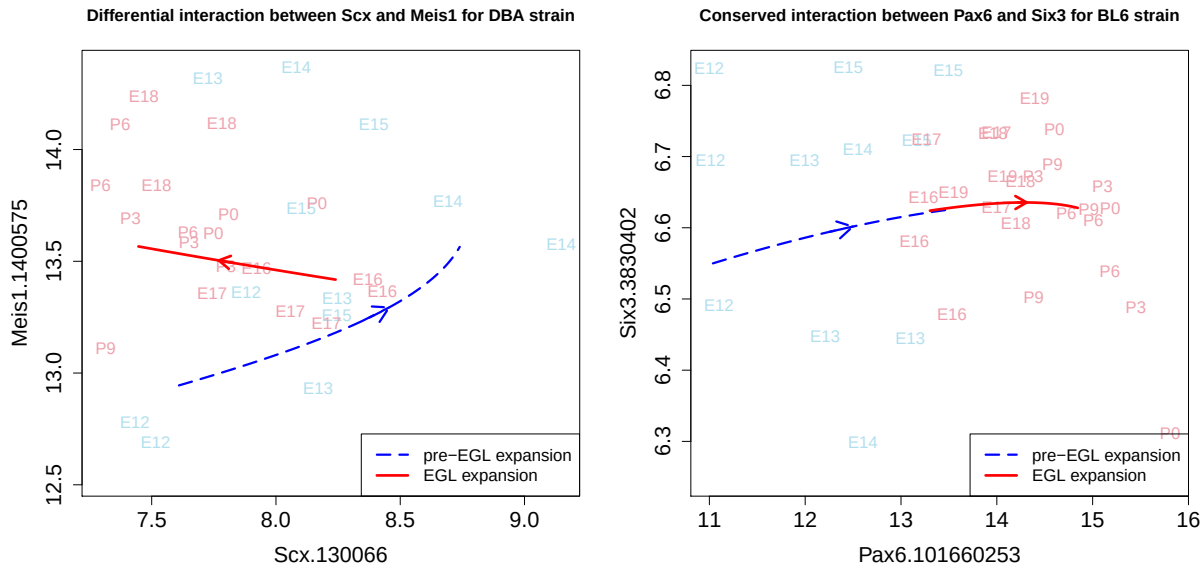

Figure S3: The phase diagrams of a differential interaction *Scx-Meis1* in the DBA strain and a conserved interaction *Pax6-Six3* in the BL6 strain, across pre-EGL stage and EGL expansion. The text in each light colored label in the plot represents the day of observation and its coordinates indicate the expression levels of corresponding genes. Arrowed lines represent trajectories computed from the homogeneous and heterogeneous models generated during CDSM fitting to the observations. Each arrowed line shows a transition in expression level over time indicated by the direction of the arrow. A blue dashed arrowed line represents a trajectory in pre-EGL stage. A red solid arrowed line represents a trajectory during EGL expansion. Left: The trajectories of *Scx-Meis1* show diverging trends between the two stages, demonstrating a differential interaction. Right: The trajectories of *Pax6-Six3* show a converging trend, demonstrating a conserved interaction between the two stages.

Child transcript name	Parent transcript name	Total interaction strength p_t	Homogeneity p_c	Heterogeneity p_d	Biological process	Brian/Neuron development
Brca1.104150463	Creb1.2230358	<0.001	<0.001	<0.001	AC	
Brca1.104150463	Creb1.1500717	0.001	<0.001	0.007	AC	
Brca1.4780669	Creb1.2230358	0.022	0.002	0.087	AC	
Creb1.1500717	Brca1.4780669	0.03	0.144	0.015	AC	
Creb1.2190471	Brca1.104150463	0.035	0.029	0.087	AC	
Creb1.6550601	Brca1.104150463	0.035	0.02	0.15	AC	
E2f1.5360093	E2f2.5270609	<0.001	<0.001	<0.001	AC	⇐
E2f1.5360093	E2f2.5570377	<0.001	0.003	<0.001	AC	⇐
E2f1.5360093	E2f2.7000465	<0.001	<0.001	<0.001	AC	⇐
E2f2.5270609	E2f1.5360093	<0.001	<0.001	<0.001	AC	
E2f2.5570377	E2f1.5360093	0.001	0.185	<0.001	AC	
E2f2.7000465	E2f1.5360093	<0.001	<0.001	0.004	AC	
Fos.1850315	Jun.840170	<0.001	<0.001	0.128	C	
Mef2c.106180288	Myc.380541	<0.001	0.042	<0.001	ACD	⇐
Mef2c.106180288	Myc.4670170	0.03	0.106	0.022	ACD	⇐
Mef2c.106980372	Myc.380541	<0.001	<0.001	0.115	ACD	⇐
Mef2c.780338	Myc.4670170	0.042	0.001	0.183	ACD	⇐
Meis1.100110440	Scx.130066	0.01	0.009	0.081	CDM	⇐
Meis1.1400575	Scx.130066	<0.001	0.012	0.008	CDM	⇐
Msx2.102190592	Pitx2.2690139	0.009	0.084	0.007	ACDM	
Msx2.102190592	Pitx2.106400039	0.012	0.148	0.008	ACDM	
Msx2.102190592	Pitx2.870537	0.012	0.045	0.019	ACDM	
Msx2.1570022	Pitx2.106400039	0.002	0.045	0.003	ACDM	
Msx2.1570022	Pitx2.870537	<0.001	0.004	0.01	ACDM	
Msx2.1570022	Pitx2.2690139	0.002	0.031	0.01	ACDM	
Myc.380541	Mef2c.670025	0.023	0.033	0.042	ACM	
Myc.380541	Pax2.7000133	0.05	0.018	0.074	ACM	
Myc.380541	Pax2.6040270	0.043	0.029	0.078	ACM	
Myc.380541	Mef2c.780338	0.012	<0.001	0.16	ACM	
Myc.380541	Mef2c.106980372	<0.001	<0.001	0.179	ACM	
Myc.4670170	Mef2c.670025	0.022	0.134	0.013	ACM	
Myc.4670170	Mef2c.106180288	0.025	0.088	0.016	ACM	
Myc.4670170	Pax2.6040270	0.047	0.069	0.056	ACM	
Pax2.7000133	Myc.380541	0.011	0.001	0.033	ACDM	
Pax3.50551	Sox10.6200538	0.035	<0.001	0.158	CD	⇐
Pax6.101660253	Six3.105130390	0.002	<0.001	0.176	CD	⇐
Pax6.101660253	Scx.130066	<0.001	<0.001	0.189	CD	⇐
Pax6.101660253	Six3.3830402	0.004	<0.001	0.198	CD	⇐
Pax6.102340114	Scx.130066	<0.001	<0.001	0.01	CD	⇐
Pax6.102340114	Six3.105130390	0.003	<0.001	0.105	CD	⇐
Pax6.102340114	Six3.3830402	0.004	<0.001	0.136	CD	⇐
Pax6.105720411	Scx.130066	<0.001	<0.001	0.173	CD	⇐
Pax6.1190025	Scx.130066	<0.001	0.003	0.002	CD	⇐
Pitx2.870537	Msx2.102190592	0.022	0.067	0.023	CDM	⇐
Scx.130066	Pax6.1190025	<0.001	0.025	<0.001		
Scx.130066	Meis1.104590215	0.001	0.067	<0.001		
Scx.130066	Meis1.1400575	<0.001	0.026	0.001		
Scx.130066	Meis1.101690520	0.002	0.172	0.001		
Scx.130066	Meis1.102190563	0.008	0.151	0.001		
Scx.130066	Pax6.102340114	<0.001	0.001	0.008		
Scx.130066	Meis1.100110440	0.007	0.009	0.016		
Scx.130066	Pax6.101660253	0.002	<0.001	0.19		
Scx.130066	Pax6.105720411	<0.001	<0.001	0.203		
Six3.105130390	Pax6.101660253	0.005	<0.001	0.193	C	⇐
Six3.3830402	Pax6.101660253	0.05	0.112	0.023	C	⇐
Sox10.6200538	Pax3.50551	<0.001	0.002	0.023	CDM	⇐
Sp3.104210341	Sp4.3850176	0.002	0.003	0.028	CDM	
Sp4.3850176	Sp3.104570037	0.01	0.003	0.092		

Table S2: A total of 58 genetic interactions are significant for the DBA strain, reported by CDSM. These were detected from 104 genetic interaction candidates experimentally verified in various organisms from BioGRID. Each transcript name is defined as <gene name>.<prob ID>, where prob ID differentiates probe locations on a microarray. We set the test size $\alpha = 0.05$. And p-values were obtained by a permutation test. An interaction is active if $p_t \leq 0.05$. All 58 interactions are active. An interaction is differential if $p_d \leq 0.05$, marked in light blue shading, and 36 interactions are so. An interaction is conserved if $p_c \leq 0.05$ and $p_d > 0.05$, marked in light yellow shading, and 21 interactions are so. According to the gene ontology of each child gene, an interaction implicated in **A**poptosis, **C**ell cycle or proliferation, **D**ifferentiation and **M**orphogenesis, is marked by A, C, D, or M in the biological process column. And each interaction known to be involved in brain/neuron development is marked by ⇐.

Child transcript name	Parent transcript name	Total interaction strength p_t	Homogeneity p_c	Heterogeneity p_d	Biological process	Brian/Neuron development
Brca1.104150463	Creb1.2190471	0.011	0.01	0.095	AC	
Brca1.4780669	Creb1.1500717	0.037	0.016	0.065	AC	
Brca1.4780669	Creb1.2190471	0.033	0.008	0.068	AC	
Brca1.4780669	Creb1.2230358	<0.001	<0.001	0.021	AC	
Brca1.4780669	Creb1.3610600	0.001	0.004	0.009	AC	
Brca1.4780669	Creb1.6550601	0.007	0.015	0.015	AC	
Creb1.1500717	Brca1.4780669	0.016	0.072	0.018	AC	
Creb1.2230358	Brca1.104150463	<0.001	<0.001	0.008	AC	
Creb1.2230358	Brca1.4780669	<0.001	<0.001	0.15	AC	
Creb1.3610600	Brca1.104150463	<0.001	0.033	0.001	AC	
Creb1.3610600	Brca1.4780669	<0.001	0.021	0.007	AC	
E2f1.5360093	E2f2.5570377	<0.001	0.008	0.009	AC	⇐
E2f2.5570377	E2f1.5360093	<0.001	0.079	<0.001	AC	
Fos.1850315	Jun.840170	0.011	0.001	0.229	C	
Jun.840170	Fos.1850315	0.002	<0.001	0.23	ACDM	⇐
Mef2c.106180288	Myc.380541	0.03	0.055	0.034	ACD	⇐
Mef2c.106180288	Myc.4670170	0.044	0.063	0.051	ACD	⇐
Mef2c.670025	Myc.380541	<0.001	0.152	<0.001	ACD	⇐
Mef2c.670025	Myc.4670170	<0.001	0.132	<0.001	ACD	⇐
Mef2c.780338	Myc.380541	0.009	<0.001	0.188	ACD	⇐
Meis1.100110440	Mef2c.106180288	<0.001	0.139	<0.001	CDM	⇐
Meis1.100110440	Msx2.1570022	<0.001	0.037	<0.001	CDM	⇐
Meis1.100110440	Scx.130066	<0.001	0.036	<0.001	CDM	⇐
Msx2.102190592	Pitx2.870537	0.014	0.008	0.09	ACDM	
Msx2.1570022	Pitx2.870537	0.049	0.047	0.054	ACDM	
Myc.380541	Mef2c.106980372	<0.001	0.183	<0.001	ACM	
Myc.380541	Mef2c.670025	0.001	0.208	0.001	ACM	
Myc.380541	Mef2c.780338	0.003	<0.001	0.196	ACM	
Myc.380541	Pax2.6040270	0.001	0.121	0.002	ACM	
Myc.380541	Pax2.7000133	<0.001	0.115	<0.001	ACM	
Myc.4670170	Pax2.7000133	0.023	0.037	0.037	ACM	
Pax2.6040270	Myc.380541	<0.001	0.087	<0.001	ACDM	
Pax2.7000133	Myc.380541	0.019	0.005	0.114	ACDM	
Pax2.7000133	Myc.4670170	0.029	0.042	0.04	ACDM	
Pax6.101660253	Scx.130066	<0.001	<0.001	0.215	CD	⇐
Pax6.101660253	Six3.105130390	<0.001	<0.001	0.216	CD	⇐
Pax6.101660253	Six3.3830402	<0.001	<0.001	0.191	CD	⇐
Pax6.105720411	Scx.130066	<0.001	<0.001	0.162	CD	⇐
Pax6.105720411	Six3.105130390	<0.001	<0.001	0.166	CD	⇐
Pax6.105720411	Six3.3830402	<0.001	<0.001	0.169	CD	⇐
Pax6.1190025	Scx.130066	<0.001	0.007	0.008	CD	⇐
Pax6.1190025	Six3.105130390	<0.001	0.004	0.003	CD	⇐
Pitx2.106400039	Msx2.102190592	0.049	0.117	0.029	CDM	⇐
Pitx2.2690139	Msx2.102190592	0.019	0.064	0.012	CDM	⇐
Scx.130066	Meis1.104590215	0.036	<0.001	0.188		
Scx.130066	Pax6.101660253	<0.001	<0.001	0.194		
Scx.130066	Pax6.1190025	0.004	0.063	0.005		
Six3.3830402	Pax6.101660253	<0.001	<0.001	0.184	C	⇐
Sox10.6200538	Pax3.50551	<0.001	<0.001	0.182	CDM	⇐
Sp4.3850176	Sp3.104210341	<0.001	0.004	0.004		
Sp4.3850176	Sp3.104570037	<0.001	0.008	0.007		
Sp4.3850176	Sp3.3840338	<0.001	0.009	<0.001		

Table S3: A total of 52 significant genetic interactions for the BL6 strain reported by CDSM. These were detected from 104 genetic interaction candidates experimentally verified in various organisms from BioGRID. Each transcript name is defined as <gene name>.<prob ID>, where prob ID differentiates probe locations on a microarray. We set the test size $\alpha = 0.05$. And p-values were obtained by a permutation test. All 52 interactions are active ($p_t \leq 0.05$). We found 30 differential interactions ($p_d \leq 0.05$), marked in light blue shading, and 21 conserved interactions ($p_c \leq 0.05$ and $p_d > 0.05$), marked in light yellow shading. According to the gene ontology of each child gene, an interaction implicated in **A**poptosis, **C**ell cycle or proliferation, **D**ifferentiation and **M**orphogenesis, is marked by A, C, D, or M in the biological process column. And each interaction known to be involved in brain/neuron development is marked by ⇐.

activation of *E2f1* and *E2f2*. Although the functions of many *E2f* proteins (*E2f1-3*) seem highly similar, it has been reported that the ability to robustly induce apoptosis is limited to *E2f1* (DeGregori *et al.*, 1997; Kowalik *et al.*, 1998). Since EGL expansion is characterized by heavy proliferation of granule cell precursors, it may be that the interaction between *E2f1* and *E2f2* is altered to effectively silence or down-regulate *E2f1* activity.

Creb-Brca1. It is known that *Creb* directly regulates *Brca1* (Atlas *et al.*, 2001). Ghosh *et al.* (2008) used ChIP methods to demonstrate that *Creb* directly interacts with the *Brca1* promotor in cell lines (Ghosh *et al.*, 2008). *Creb* has been demonstrated to enhance neuronal survival and regulate apoptosis in cerebellar granule cells (Bonni *et al.*, 1999; Mantamadiotis *et al.*, 2002). In contrast, the tumor suppressor gene, *Brca1* is robustly expressed by proliferating granule cell precursors in the EGL (Korhonen *et al.*, 2003). Thus, it is plausible that *Brca1* and *Creb1* interact to regulate cell proliferation in the EGL expansion phase. This putative interaction could produce a differential interaction between these two genes, when compared to the pre-EGL stage.

Pax6-Six3. They have been shown to interact during eye development, and this interaction is necessary for the proper development of the structure. Experiments indicate the presence of mutual binding sequences upstream for each of the transcription factors (Goudreau *et al.*, 2002). In the cerebellum, nothing is known about the relationship between *Pax6* and *Six3*, however expression between *Pax6* and *Six3* is coincident in the EGL at P0 (Engelkamp *et al.*, 1999; Conte *et al.*, 2005).

Pax6. *Pax6* is involved in regulating neuronal migration, morphology and proliferation in cerebellar granule cells (Swanson *et al.*, 2005; Duparc *et al.*, 2006). The pre-EGL stage is characterized by neuronal migration of granule cell precursors into the EGL, and the post-EGL stage is characterized by proliferation of granule cells, followed by differentiation and inward migration. *Pax6* is robustly expressed in granule neurons throughout both temporal periods (Engelkamp *et al.*, 1999) and could therefore regulate migration or proliferation.

Msx2. *Msx2* has been reported to be important in the formation of the cerebellum since *msx2-/-* mice show abnormalities in cerebellar formation (Satokata *et al.*, 2000).

Pitx2. *Pitx2* has recognized roles in ocular (eye), tooth, and pituitary gland development. However, in at least one case, mutation of this gene has been associated with cerebellar malformations (Idrees *et al.*, 2006). In another study in mice, partial loss of *Pitx2* (*Pitx2-/-* is embryonic lethal) initially leads to no obvious defects in brain development up to stage E10.5, whereas midbrain defects are visible during later stages (E14.5) (Martin *et al.*, 2004). *Msx2* and *Pitx2* have been noted to interact during the formation of tooth anlage in mammals in a mutually repressive manner. Specifically, these two factors compete for binding to promoter regions in genes essential for tooth development (Green *et al.*, 2001).

Scx. *Scx* is associated with several nervous system diseases in humans, and has been implicated in neurodegeneration (Yeghiazaryan *et al.*, 1999). It is possible that *Scx* interacts with other transcription factors such as *Pax6*, as predicted by our comparative modeling, during cerebellar development.

9 Exploring novel gene interactions during cerebellar development

The CDSM was also applied to discover novel genetic interactions during cerebellar development by genome-wide exploration. The genome-wide time course expression data contained more than 40,000 transcripts, on which we performed gene selecting, gene clustering, and CDSM. First, we merged time course observations of two mouse strains by a gene selection process. The transcripts with consistent expression between DBA and BL6 mouse strains were selected in order to find comparative gene interactions that are conserved between the two strains. For the selection, we used a linear statistical model

$$EXP_{ijk} = Time_i + Strain_j + \sigma_{ijk} \tag{28}$$

where Exp_{ijk} represent the expression value of the kth transcript replicate in strain j at time i, $Time$ and $Strain$ are two independent factors in this statistical model, σ is a normally distributed noise. If $Strain$

Cluster name	a cluster member	p_t (total strength)	p_c (homogeneity)	p_d (heterogeneity)
C9	*Ppard*	1.94e-35	2.48e-18	5.05e-20
C326	*Lef1*	1.63e-54	2.14e-32	6.83e-26
C879	*Crebbp*	2.28e-28	8.78e-20	1.18e-11
C1715	*Map2k7*	5.96e-28	3.34e-17	8.14e-14
C164	*Nfatc2*	2.33e-35	7.28e-30	3.87e-09
C656	*Ccnd2*	5.00e-42	1.92e-31	5.54e-14
C1308	*Nfatc1*	4.59e-47	2.27e-28	3.11e-22
C185	*Ruvbl1*	2.55e-37	1.18e-15	1.54e-24
C1264	*Mapk8*	1.97e-30	2.55e-17	5.56e-16
C169	*Mapk8*	6.18e-46	2.13e-39	9.09e-11
C39	*Ctbp2*	6.60e-44	1.15e-28	8.09e-19

Table S4: Statistical significance of comparative modeling for clusters containing genes involved with the WNT pathway. These gene clusters were all significantly differentially influenced by their parent clusters.

factor is insignificant ($\alpha = 0.5$ means highly conserved between strains) of a transcript, it will be selected. A total of 16,239 transcripts were thus selected. Second, gene clustering was used to group those transcripts into clusters, because linearly correlated transcripts are mathematically equivalent in DSM. The Pearson correlation coefficient was utilized to measure the linearity between any two transcripts by analyzing their time course expression levels. For each of the resulting 1,823 clusters, one representative that is most linearly correlated to other members within a cluster was chosen. A DSM is used to represent expression rates of 1,823 cluster representatives, and each of them is modeled by

$$\frac{dx_i}{dt} = \beta_{i0} + \sum_{j \in Par_i} \beta_{ij} \frac{x_j^2}{8^2 + x_j^2} - \beta_i x_i$$

where x_i represents the expression level of gene i, Par_i the influence genes of gene i, βs the model coefficients. The CDSM with a maximum parent number ($dim(Par_i)$) of 2 was used, and self influence is allowed.

Putative interactions were generated by the CDSM for 1,823 cluster representatives across two development events. The two temporal periods compared were the *presence of rhombic lip* (E12-E17) *v.s.* the *presence of distinct inner EGL* (E18-P9). In this example, we illustrate a subset of interactions involving TFs and their target genes in the WNT signaling pathway. The statistical significance of each node is given by p-values in Table S4. A TF named *Lef1* involved in the WNT pathway was significantly influenced differently between the two temporal events. And genes in a cluster (C851) were very active in influencing four genes involved in the WNT pathway. Genes in cluster C851 are: *Rsc1a1*, *Zfp367*, *Rnf26*, *Gli2*, *Impa2*, *Rfc1*, *Polb*, *Fen1*, *B830045N13Rik*, *BC062185*, *Uros*, *Prr14*, *Bik*, *Zkscan17*, *Pmf1*, *Mybl2*, *Tk1*, *Tbl2*, and *Plec1*.

References

Atlas, E., Stramwasser, M., and Mueller, C. R. (2001). A CREB site in the BRCA1 proximal promoter acts as a constitutive transcriptional element. *Oncogene*, **20**(48), 7110–7114.

Bonni, A., Brunet, A., West, A. E., Datta, S. R., Takasu, M. A., and Greenberg, M. E. (1999). Cell survival promoted by the Ras-MAPK signaling pathway by transcription-dependent and -independent mechanisms. *Science*, **286**(5443), 1358–1362.

Conte, I., Morcillo, J., and Bovolenta, P. (2005). Comparative analysis of Six 3 and Six 6 distribution in the developing and adult mouse brain. *Developmental Dynamics*, **234**(3), 718–725.

DeGregori, J., Leone, G., Miron, A., Jakoi, L., and Nevins, J. R. (1997). Distinct roles for E2F proteins in cell growth control and apoptosis. *PNAS*, **94**(14), 7245–7250.

Duparc, R. H., Boutemmine, D., Champagne, M. P., Tetreault, N., and Bernier, G. (2006). Pax6 is required for delta-catenin/neurojugin expression during retinal, cerebellar and cortical development in mice. *Developmental Biology*, **300**(2), 647–655.

Engelkamp, D., Rashbass, P., Seawright, A., and Van Heyningen, V. (1999). Role of Pax6 in development of the cerebellar system. *Development*, **126**(16), 3585–96.

Ghosh, A., Shuman, S., and Lima, C. D. (2008). The structure of Fcp1, an essential RNA polymerase II CTD phosphatase. *Molecular Cell*, **32**(4), 478–490.

Goudreau, G., Petrou, P., Reneker, L. W., Graw, J., Löster, J., and Gruss, P. (2002). Mutually regulated expression of Pax6 and Six3 and its implications for the Pax6 haploinsufficient lens phenotype. *PNAS*, **99**(13), 8719–8724.

Green, P. D., Hjalt, T. A., Kirk, D. E., Sutherland, L. B., Thomas, B. L., Sharpe, P. T., Snead, M. L., Murray, J. C., Russo, A. F., and Amendt, B. A. (2001). Antagonistic regulation of Dlx2 expression by PITX2 and Msx2: implications for tooth development. *Gene Expression*, **9**(6), 265–281.

Heckman, N. and Ramsay, J. O. (1996). *Spline smoothing with model based penalties*. R package version 1.0-14 (2010). S original by Jim Ramsay. R port by Brian Ripley.

Hu, R., Qiu, X., Glazko, G., Klebanov, L., and Yakovlev, A. (2009). Detecting intergene correlation changes in microarray analysis: a new approach to gene selection. *BMC Bioinformatics*, **10**(1), 20.

Idrees, F., Bloch-Zupan, A., Free, S. L., Vaideanu, D., Thompson, P. J., Ashley, P., Brice, G., Rutland, P., Bitner-Glindzicz, M., Khaw, P. T., Fraser, S., Sisodiya, S. M., and Sowden, J. C. (2006). A novel homeobox mutation in the PITX2 gene in a family with Axenfeld-Rieger syndrome associated with brain, ocular, and dental phenotypes. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics*, **141B**(2), 184–191.

Korhonen, L., Brännvall, K., Skoglösa, Y., and Lindholm, D. (2003). Tumor suppressor gene BRCA-1 is expressed by embryonic and adult neural stem cells and involved in cell proliferation. *Journal of Neuroscience Research*, **71**(6), 769–776.

Kowalik, T. F., DeGregori, J., Leone, G., Jakoi, L., and Nevins, J. R. (1998). E2F1-specific induction of apoptosis and p53 accumulation, which is blocked by Mdm2. *Cell Growth & Differentiation*, **9**(2), 113–118.

Mantamadiotis, T., Lemberger, T., Bleckmann, S. C., Kern, H., Kretz, O., Martin Villalba, A., Tronche, F., Kellendonk, C., Gau, D., Kapfhammer, J., Otto, C., Schmid, W., and Schütz, G. (2002). Disruption of CREB function in brain leads to neurodegeneration. *Nature Genetics*, **31**(1), 47–54.

Martin, D. M., Skidmore, J. M., Philips, S. T., Vieira, C., Gage, P. J., Condie, B. G., Raphael, Y., Martinez, S., and Camper, S. A. (2004). PITX2 is required for normal development of neurons in the mouse subthalamic nucleus and midbrain. *Developmental Biology*, **267**(1), 93–108.

Oliver, T. G., Grasfeder, L. L., Carroll, A. L., Kaiser, C., Gillingham, C. L., Lin, S. M., Wickramasinghe, R., Scott, M. P., and Wechsler-Reya, R. J. (2003). Transcriptional profiling of the Sonic hedgehog response: a critical role for N-myc in proliferation of neuronal precursors. *PNAS*, **100**(12), 7331–7336.

Ouyang, Z. and Song, M. (2009). Comparative Identification of Differential Interactions from Trajectories of Dynamic Biological Networks. In *Proceedings of German Conference on Bioinformatics Halle Germany*, pages 163–172.

Prill, R. J., Marbach, D., Saez-Rodriguez, J., Sorger, P. K., Alexopoulos, L. G., Xue, X., Clarke, N. D., Altan-Bonnet, G., and Stolovitzky, G. (2010). Towards a rigorous assessment of systems biology models: the DREAM3 challenges. *PLoS ONE*, **5**(2), e9202.

Satokata, I., Ma, L., Ohshima, H., Bei, M., Woo, I., Nishizawa, K., Maeda, T., Takano, Y., Uchiyama, M., Heaney, S., Peters, H., Tang, Z., Maxson, R., and Maas, R. (2000). Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nature Genetics*, **24**(4), 391–395.

Stark, C., Breitkreutz, B.-J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006). BioGRID: a general repository for interaction datasets. *Nucleic Acids Research*, **34**(Database issue), D535–D539.

Stark, C., Breitkreutz, B.-J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M. S., Nixon, J., Van Auken, K., Wang, X., Shi, X., Reguly, T., Rust, J. M., Winter, A., Dolinski, K., and Tyers, M. (2010). The BioGRID Interaction Database: 2011 update. *Nucleic Acids Research*, pages D698–D704.

Swanson, D. J., Tong, Y., and Goldowitz, D. (2005). Disruption of cerebellar granule cell development in the Pax6 mutant, Sey mouse. *Developmental Brain Research*, **160**(2), 176–193.

Tyson, J. J. (1991). Modeling the cell division cycle: cdc2 and cyclin interactions. *PNAS*, **88**(16), 7328–7332.

Yeghiazaryan, K., Turhani-Schatzmann, D., Labudova, O., Schuller, E., Olson, E. N., Cairns, N., and Lubec, G. (1999). Downregulation of the transcription factor scleraxis in brain of patients with Down syndrome. *Journal of Neural Transmission. Supplementum*, **57**, 305–314.

Zar, J. H. (2009). *Biostatistical Analysis*. Prentice Hall, 5th edition.

Zhu, D., Li, Y., and Li, H. (2007). Multivariate correlation estimator for inferring functional relationships from replicated genome-wide data. *Bioinformatics*, **23**(17), 2298–2305.