Asymmetric robustness in the feedback control of the retinoic acid network response to environmental disturbances

Madhur Parihar[1,$], Liat Bendelac-Kapon[2,$], Michal Gur[2,$], Abha Belorkar[1], Sirisha Achanta[1], Keren Kinberg[2], Rajanikanth Vadigepalli[1,*], Abraham Fainsod[2,*]

[1]Daniel Baugh Institute for Functional Genomics/Computational Biology, Department of Pathology, Anatomy, and Cell Biology, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, Pennsylvania, 19107 USA

[2]Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel-Canada, Faculty of Medicine, Hebrew University, Jerusalem 9112102, Israel

[$]equal contribution

*Corresponding authors: Rajanikanth.Vadigepalli@jefferson.edu

Abraham.Fainsod@mail.huji.ac.il

Review Codes for online datasets GEO SuperSeries GSE154408: cryfqygubnszryt

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE154408

ABSTRACT

Robustness is a characteristic of regulatory pathways to ensure signal consistency in light of environmental changes or genetic polymorphisms. The retinoic acid (RA) pathway is a central developmental and tissue homeostasis regulatory signal, strongly dependent on nutritional sources of retinoids and affected by exogenous chemicals. We performed transient physiological RA signaling disturbances during embryogenesis followed by kinetic transcriptomic and high-throughput qPCR analysis of the recovery. Unbiased pattern analysis identified the RA metabolic network as the main regulated module aimed at achieving signaling robustness. We used a principal trajectory-based analysis of the clutch-dependent variability and organized the results into a robustness efficiency matrix comparing the RA feedback regulation and *hox* gene expression (RA targets). We found the feedback autoregulation to be sensitive to the direction of the RA perturbation: RA knockdown exhibited an upper response threshold, whereas RA addition did not activate a feedback response below a minimum threshold. These results demonstrate an asymmetric capacity for robust feedback control of the RA signal during early embryogenesis, probably based on genetic polymorphisms, likely a significant contributor to the manifestation of developmental defects.

KEYWORDS

Embryo development; Retinoic acid; Xenopus embryo; time-series transcriptomics; temporal gene expression pattern analysis; developmental trajectory analysis; autoregulatory feedback control.

Introduction

Retinoic acid (RA) is a central regulatory signaling pathway active during embryogenesis and adult tissue homeostasis (le Maire and Bourguet, 2014; Metzler and Sandell, 2016). RA is synthesized from vitamin A (retinol) or other retinoids or carotenoids obtained from the diet (Blaner et al., 2016; Fainsod and Kot-Leibovich, 2018; Ghyselinck and Duester, 2019; Kedishvili, 2016). During embryogenesis, excessive or reduced RA signaling induces a wide array of embryonic malformations (Corcoran et al., 2002; Durston et al., 1989; Hollemann et al., 1998; Kessel and Gruss, 1991; Kot-Leibovich and Fainsod, 2009; Marshall et al., 1992; Papalopulu et al., 1991). Disorders linked to reduced RA signaling include vitamin A deficiency syndrome, DiGeorge/VeloCardioFacial syndrome, Fetal Alcohol Spectrum Disorder, Congenital Heart Disease, neural tube defects, and multiple types of cancer (Hartomo et al., 2015; Kim et al., 2005; Kot-Leibovich and Fainsod, 2009; Pangilinan et al., 2014; See et al., 2008; Urbizu et al., 2013). Therefore, RA levels are tightly regulated throughout life, controlling quantitative, spatial, and temporal expression and activity of RA biosynthetic and metabolizing enzymes (Blaner, 2019; Ghyselinck and Duester, 2019; Kedishvili, 2016).

RA biosynthesis involves two sequential oxidation steps: first, mainly alcohol dehydrogenases (ADH) or short-chain dehydrogenase/reductases (SDR) oxidize vitamin A (retinol, ROL) to retinaldehyde (RAL), followed by the retinaldehyde dehydrogenase (RALDH) catalyzed oxidation of RAL to RA (Ghyselinck and Duester, 2019; Kedishvili, 2016; Metzler and Sandell, 2016). ROL and RAL can also be produced from retinyl esters or ß-carotene from food sources (Blaner, 2019). RA availability is further affected by ROL, RAL, and RA binding proteins (Napoli, 2017). In gastrula vertebrate embryos, RA signaling is triggered by the expression of RALDH2 (ALDH1A2) completing the biosynthesis of RA (Begemann et al., 2001; Chen et al.,

2001; Grandel et al., 2002; Niederreither et al., 1999). Substrate (RAL) availability for the RALDH enzymes is controlled by RAL reductases (Adams et al., 2014; Billings et al., 2013; Feng et al., 2010; Porté et al., 2013; Shabtai and Fainsod, 2018). Besides the temporal and spatial regulation of RA metabolic network components, other RA network components, including the RA receptors (RAR and RXR) and retinoid-binding proteins (Cui et al., 2003; Janesick et al., 2015; Lohnes et al., 1995) are under tight regulation. Then, RA signaling is controlled through its regulated biosynthesis and degradation, as well as restricted spatiotemporal expression of the metabolic and gene-regulatory components, which together are expected to provide robustness to RA signaling. The maternal nutritional status and environmental exposure to chemicals such as alcoholic beverages (ethanol) and others can also affect the RA signal (Paganelli et al., 2010; Shabtai et al., 2018). These observations point to the close interaction of RA signaling and the environment and the necessity to adapt the RA signaling network to nutritional changes and insults. This adaptation and maintenance of normal signaling levels under changing conditions is termed robustness (Eldar et al., 2004; Nijhout et al., 2019).

A deeper understanding of the RA signaling pathway during embryogenesis is required to elucidate its multiple regulatory roles and the regulation of the signaling robustness in the presence of disturbances. Commonly, the RA pathway is studied by increasing this signal (Durston et al., 1989; Kessel, 1992; Sive et al., 1990), or knockdown studies take advantage of inhibitors, inverse agonists, or enzymatic degradation (Hollemann et al., 1998; Janesick et al., 2013; Janesick et al., 2014; Kot-Leibovich and Fainsod, 2009). In multiple RA knockdown studies, the developmental malformations observed were milder than expected (Blumberg et al., 1997; Hollemann et al., 1998; Janesick et al., 2014; Koide et al., 2001; Shabtai et al., 2018; Sharpe and Goldstone, 1997) suggesting a compensatory mechanism. To characterize the

robustness of RA signaling during early embryogenesis, we devised clearly defined and transient RA manipulations in the physiological range that are terminated during early gastrula, and monitored by RNAseq, qPCR, and phenotypic analysis for several hours during recovery. These results demonstrated high RA signaling robustness resulting in relatively small transcriptome-wide changes. Unbiased detailed transcriptomic analysis revealed that RA metabolic and signaling network components exhibited expression changes dependent on the direction of RA manipulation, suggesting a mechanistic explanation for the robustness observed. Analysis of multiple embryo clutches (biological replicates) revealed differences in their individual robustness response to enhanced or reduced RA signaling. These results exposed potential consequences of the underlying genetic differences between clutches. Our results suggest an asymmetric capacity for RA signaling robustness in the early embryo, likely contributing to the human developmental defects that arise due to RA imbalance, most often, reduced Vitamin A levels during early pregnancy.

Results

Physiological RA manipulation uncovers signaling robustness

To manipulate RA signaling in early embryos, we routinely employ two approaches to knockdown this ligand (Fig. 1A). We either partially inhibit the oxidation of RAL to RA with RALDH inhibitors like 4-diethylaminobenzaldehyde (DEAB) (Shabtai et al., 2016), or we overexpress the RA hydroxylase, CYP26A1 (Hollemann et al., 1998; Yelin et al., 2005), to render RA inactive (Topletz et al., 2015). These treatments result in unexpectedly mild developmental defects (Figs. 1B-D). To induce severe phenotypes by RA loss-of-function we combine knockdown treatments targeting the metabolism at different steps (Fig. 1A, E) supporting the robustness of RA signaling.

To study the RA network robustness in the physiological range we empirically determined the RA concentrations to use by testing concentrations (1 nM - 1 μM) spanning the reported physiological content of *X. laevis* early embryos (100-150 nM all-*trans* RA) (Durston et al., 1989; Kraft and Juchau, 1992; Kraft et al., 1995; Sive et al., 1990). We analyzed gene expression of RA metabolic enzymes and RA targets during early (st. 10.25) and late gastrula (st.12) (Nieuwkoop and Faber, 1967). Genes positively regulated by RA at both stages, *hoxb1*, *cyp26a1*, and *dhrs3* exhibited dose-dependent responses (Fig. 2A-C). The genes *raldh2* and *rdh10,* are downregulated by most RA doses (Fig. 2D, F), while *raldh3* (*aldh1a3*) is upregulated earlier on and is repressed later on (Fig. 2E). Importantly, RA concentrations as low as 10 nM had a significant effect on gene expression.

We performed transient pharmacological RA manipulations within the physiological range in early *Xenopus* embryos. The perturbations were terminated by washing the treatments and then the embryos were monitored during the recovery period. Embryos were treated with RA (10 nM) for 2 hours starting from late blastula (st. 9.5) and washed during early gastrula (st. 10.25). At different time points during the post-wash recovery period, samples were collected to perform kinetic analysis of the changes in gene expression by qPCR (Fig. 3). To monitor the changes in RA signaling levels by gene expression, two well-characterized RA-regulated genes, *hoxb1* and *hoxb4,* were analyzed. Both genes showed that at the time of treatment washing (t_0), they were upregulated when compared to control sibling embryos (Fig. 3A). Two hours into the recovery, both RA target genes were back to control expression levels. To begin understanding the robustness of RA signaling and the dynamic regulation of RA metabolism we studied the expression of *dhrs3*, which preferentially reduces RAL to ROL to attenuate RA biosynthesis (Feng et al., 2010), *cyp26a1*, which targets RA for degradation, and *raldh2* that produces RA from retinaldehyde (Shabtai et al.,

2016). Both negative RA regulators, *dhrs3* and *cyp26a1* are upregulated at t_0 by increased RA (Fig. 3B) and return to normal expression levels after 4-hour recovery. As expected, *raldh2* exhibited an RA-promoted downregulation at t_0 (Fig. 3B), returning to normal levels after 4 hours. These results indicate that the gene expression of RA metabolic enzymes, auto-regulated by RA (Fig. 2), is shifted to achieve normal RA levels in the face of external RA perturbation.

A complementary study was performed by inhibiting RA biosynthesis with DEAB (Fig. 3C,D). All genes studied exhibited fluctuations during the recovery period. Also, in this case, *hoxd1* and *hoxb1* reached almost normal levels at an earlier stage than the genes encoding RA metabolic components (Fig. 3C). Genes encoding anabolic enzymes, e.g., *raldh2*, were upregulated, and catabolic components, e.g., *cyp26a1*, were downregulated (Fig. 3D).

Physiological RA perturbations uncover signaling robustness at the transcriptomic scale

To gain a better understanding of the RA signaling robustness, we performed a kinetic transcriptomic analysis. To optimize the kinetic study, parameters such as developmental stages and timing between samples were empirically tested. We treated the embryos with RA (10 nM) or DEAB (50 μM) for 2 hours from late blastula (st. 9.5) to early gastrula (st. 10.25) and collected samples every 1.5 hours after terminating the treatment (Fig. 4A). For each biological repeat, all treatments, controls, and time points were collected from a single clutch (eggs) of a single female. The efficiency of the treatments and quality of the RNA samples was first monitored by qPCR for changes in *hox* gene expression as a readout of the RA signal levels. Only those biological repeats where both treatments exhibited the expected upregulation (RA) or downregulation (DEAB) in *hox* gene expression were selected for RNA-seq.

We analyzed the time series transcriptomic data set for differentially expressed genes using a two-way ANOVA (t=0, 1.5, 3, 4.5 hours; treatments: RA, DEAB, control; n=6 biological replicates). Principal Component Analysis (PCA) of the gene expression variation showed separation along the first principal component (PC1) which corresponds to the developmental stage (Fig. 4B). Surprisingly, RA manipulated samples clustered with the control samples of the same developmental stage. Normal transcriptomic changes as a result of progression through embryogenesis appear to be the dominant variable distinguishing the samples irrespective of treatment (Fig. 4B). The second component (PC2) separated the sample groups at intermediate time points (1.5 and 3h) from those at 0 and 4.5h time points, indicating a transient differential expression shift as the next most dominant pattern in the data. The effects of RA and DEAB on the transcriptome was not readily apparent in the next seven principal components. The PC8 showed some separation of the DEAB group from the RA and Control groups (Fig. 4C), whereas PC10 showed some separation of the RA treatment group from Control and DEAB treatments (Fig. 4D). The top-ranked genes along PC8 and PC10 showed distinct dynamic patterns across the treatments, whereas top-ranked genes along PC1 and PC2 largely corresponded to in-common dynamic changes over time (Fig. 4D; Supplemental Fig. S1). The overall magnitude of induced changes in gene expression in response to RA or DEAB appears to be less than the normal transcriptome changes occurring during early developmental stages, showing that RA signaling robustness possibly dampens the gene expression changes otherwise induced by abnormal RA levels. These results are consistent with the response of a robust system that functions to limit the gene expression changes in the majority of the transcriptome.

Dynamic pattern analysis reveals the molecular genetic mechanisms underlying the RA robustness response

For characterization of the molecular transcriptomic mechanism of the RA robustness response, we analyzed the clutch-averaged data using an unbiased dynamic pattern analysis approach to categorize the gene expression profiles along all possible discretized patterns (Kuttippurathu et al., 2016). Our analysis revealed a total of 4693 significantly differentially expressed genes with greater than 2-fold changes over time (compared to t_0 within each treatment group; multiple testing corrected q-value < 0.05). For each gene, a pattern based on the direction of expression change above a 2-fold threshold, or no change at each of the three recovery time points relative to t_0 was determined. Such an exhaustive approach allows us to enumerate the dominant as well as subtle patterns in the data overcoming limitations of conventional cluster analysis that could miss or mask the smaller groups of genes with distinct expression profiles over time. Out of the 27 (3*3*3) possible dynamic patterns, only 10 patterns were exhibited by genes in the transcriptome in at least one of the three experimental groups (RA, DEAB, Control) (Fig. 5A). Of these, only 5 patterns were exhibited by a substantial number of genes (>100) in any experimental group, corresponding to up- or downregulation at later time points (3h and 4.5h) (Fig. 5A). At the 2-fold threshold, there were no genes that showed downregulation first and then upregulation and vice versa. There were no significant differences in the gene counts between RA or DEAB and control groups (two-tailed Z test, $p > 0.05$). Gene Ontology enrichment analysis of the control group highlighted several functional annotations including development, cell cycle, morphogenesis, RA biosynthesis, gastrulation, and others. (Fig. 5B; Supplemental Table S1). The patterns with relatively fewer genes (<100) correspond to transient up- or downregulation at 3h that is normalized by 4.5h, as well as progressive downregulation.

The pattern analysis (Fig. 5) shows that at the 2-fold threshold, there are no distinctive dynamic gene expression patterns that occur only in response to the RA or DEAB treatments. Also, the number of genes changing per pattern is similar between the two treatments and control suggesting that the physiologically relevant concentrations of RA and DEAB modulate the RA network components at a scale resembling the changes in controls. For instance, only a subset of genes that showed late downregulation (first row, Fig. 5A) was common across all the three groups (772 genes, Fig. 5C). Several genes showed similar late downregulation between the treatments or between one of the treatments and the Control group (Fig. 5C). In addition, many genes showed a late downregulation pattern only in one of the samples (Fig. 5C). While the Venn diagram analysis (Fig. 5C) revealed extensive overlap between both treatments and Control samples for each pattern, the overlap across patterns is not immediately clear from this analysis.

To exhaustively compare the RA and DEAB treatment groups for their effect on gene expression, we adapted our unbiased approach named COMPACT for analyzing time-series differential transcriptomic profiles across multiple experimental conditions (Kuttippurathu et al., 2016). For statistically significant changing genes within each treatment group (two-way ANOVA; $q<0.05$), we constructed discrete patterns based on differential gene expression between treatment (RA or DEAB) and controls (81 theoretically possible patterns, with up-, down- and no-regulation at t=0, 1.5, 3 and 4.5 h). Only the patterns with at least one gene in either comparison were included, yielding 32 distinct patterns (Fig. 6A). Finally, we intersected the two distributions to create a matrix of comparative patterns where each element corresponds to a distinct pair of patterns corresponding to RA vs. control and DEAB vs. control (Fig. 6B). The discrete patterns were based on a 1.3-fold average difference between RA (or DEAB) and control groups at each of the four time points. The effect

size threshold was chosen at a lower level than 2-fold as the differential expression analysis revealed that the RA/DEAB perturbations were leading to a smaller magnitude of changes at each time point, as compared to the larger changes occurring normally over time, additional suggestive evidence of robustness.

The COMPACT matrix (Fig. 6B, Extended Data 1) shows that the majority of the genes sensitive to RA manipulation exhibited sensitivity to only one direction of the RA perturbation. A total of 193 genes only responded to the addition of RA (middle row), whereas 224 genes were only responsive to RA knockdown (DEAB; middle column). Of the genes responsive to both RA and DEAB perturbations, 88 genes showed up- or downregulation irrespective of RA or DEAB treatment (quadrants a and d; Fig. 6B). A set of 48 genes showed opposite transcriptional outcomes to RA increase versus RA decrease (quadrants b and c; Fig. 6B). Interestingly, the set of genes that showed upregulation by exogenous RA addition and downregulation by DEAB (quadrant b) contained multiple genes involved in RA metabolism (Fig. 6C).

We mapped the expression changes in Fig. 6B to the RA signaling and metabolic network (Fig. 6D,E). Notably, genes encoding proteins involved in suppressing RA levels (*cyp26a1*, *dhrs3*) were upregulated and the genes encoding enzymes for RA production (*aldh1a2*, *aldh1a3*, and *rdh10*) were downregulated in response to transient RA increase. These genes showed opposite regulation in response to DEAB treatment (Fig. 6E). As an independent evaluation, we analyzed the differential expression time series data using another unbiased approach, weighted gene correlation network analysis (WGCNA) (Langfelder and Horvath, 2008). A WGCNA gene expression module with the highest correlation with treatment contained a similar gene set as shown in Fig. 6C, supporting the COMPACT analysis (green module, Supplemental Table S2; Extended Data 2). Taken together, our results

support a mechanism to maintain RA signaling homeostasis to counter the effects of exogenous perturbations and prevent teratogenic effects by a transcriptional feedback control system.

Clutch-wise heterogeneity uncovers alternative mechanisms for RA robustness

The analysis above shows changes in RA network components and identified the main RA network components responding during early gastrulation (Fig. 6). Interestingly, the PCA analysis suggested clutch-specific heterogeneity in gene expression changes where all samples from the same clutch tended to cluster together but the different clutches separated slightly from each other (Fig. 4B,C; Supplemental Fig. S2). For these reasons, we analyzed the transcriptional changes in the RA metabolic network within each clutch. We generated an additional set of samples to test the dynamic expression changes in RA metabolic network genes using high-throughput real-time PCR (HT-qPCR; Fluidigm Biomark; Supplemental Table S2) to validate RNAseq results. These six additional clutches (biological repeats labeled G-L) were treated with RA or DEAB following the same experimental design (Fig. 4A) used for the RNAseq study (clutches A-F). We observed a similar heterogeneity between clutches by HT-qPCR to the clutch-to-clutch variation in the RNAseq data (Supplemental Figs. S2 and S3).

We ordered the 12 biological repeats (clutches) according to the earliest time at which *hoxa1* returned to the control levels (Fig. 7) intermingling the clutches of both studies. We observed high variability in the response to the RA manipulation among the individual clutches. Clutches with significant *hoxa1* upregulation following RA addition showed limited downregulation in response to RA biosynthesis inhibition (DEAB; clutches C, L, J, K, H, I in Fig. 7). Clutches E, D, F, G, A, B (Fig. 7) showed the opposite response, i.e. robust *hoxa1* change following RA knockdown, while the RA

increase induced a mild response. The suppressors of RA signaling, e.g., *cyp26a1* and *dhrs3*, showed significantly altered expression in clutches (C, L, J, K, H) with larger deviations in *hoxa1* expression and a strong and extended response to the addition of RA. In contrast, the differential expression RA producers (e.g., *aldh1a2* and *rdh10*) after RA manipulation was relatively mild and did not fully align with the deviation in *hoxa1* expression (Fig. 7). Such a heterogeneous correlation between the extent of feedback regulation and clutch-to-clutch variation in *hoxa1* expression was observed across the RA metabolic network (Supplementary Fig. S4).

Asymmetric robustness in response to increased versus decreased RA signaling

We sought to quantitatively rank the robustness of each clutch in an integrated manner based on the expression shift of multiple *hox* genes and the feedback response of the RA metabolic network genes. We pursued a trajectory-based approach in which the samples of all biological repeats (clutches) at all time points were projected onto the first three principal components based on the expression of *hox* genes (*hoxa1*, *hoxa3*, *hoxb1*, *hoxb4*, *hoxd4*) or the RA metabolic network genes. A principal curve was fit to the projected data, representing the trajectory in which the system evolves over time for all clutches combined. Each clutch was visualized separately along the principal trajectory, allowing us to compare the temporal evolution of deviations between treated samples and controls (Figs. 8A,B; Supplemental Fig. S5). As a multi-gene measure of the robustness of each clutch, the net absolute distance between the treated samples and controls was computed at all time points along the principal curve. Our results revealed a wide range of clutch-to-clutch variability in robustness as assessed by the integrated deviation from the control trajectory of multiple *hox* genes (Fig. 8C; Supplemental Fig. S5). The distance between treated samples and controls decreased over time in nearly all the clutches with significant differences in the time taken to

close the gap (Fig. 8C). Clutches with low RA increase robustness showed larger distances between samples and controls along the *hox* gene trajectory at any time point (Fig. 8A, clutch C), compared to robust clutches with decreasing distances (Fig. 8A, clutch E). By contrast, DEAB treatment resulted in an opposite robustness pattern (Fig. 8B), with clutch C showing the least deviation and clutch E the widest deviation (Fig. 8B top row and Fig. 8C). Clutch J exhibited intermediate robustness responses to both the increase and the reduction of RA levels (Fig. 8A,B). The anti-correlated pattern of RA robustness suggests an asymmetric tradeoff in the gastrula embryo to counter changes in RA levels (Fig. 8E). The *hox* responses to changed RA levels revealed that all biological repeats aligned along a negative-slope diagonal and covered the whole range of responses (Fig. 8E). This result suggested the establishment of an RA robustness gradient among the clutches.

A similar analysis of the RA network component changes showed a scattered pattern among the biological repeats with no clear trend. The feedback regulatory response based on the RA metabolic network trajectory map (Figs. 8A,B bottom rows; Supplemental Fig. S5) was variable across clutches, with a characteristic pattern of reduced deviation over time in several clutches (Fig. 8D). There was no significant correlation in the net shift in feedback regulatory action between RA addition versus reduction (Fig. 8F). Interestingly, the scattering distribution depended on the direction of the RA manipulation. The range of clutch distribution for DEAB treatment suggests an immediate response to any RA reduction and an upper limit to this response (Fig. 8F). In contrast, the clutch distribution RA addition suggests that the response requires a threshold change to be activated and it exhibits also an upper threshold. These observations suggest a high sensitivity to RA reduction but a lower sensitivity to slight RA increase.

Clutch robustness variability affects the efficiency-efficacy of the response

We examined whether the asymmetric robustness to RA change direction correlates to the extent of the feedback regulatory effect (Fig. 9). We formulated an "efficiency-efficacy matrix" (Fig. 9A). The sections of this matrix delineate distinct possibilities based on all permutations of the robustness level based on the net shift in RA metabolic network, and the phenotypic outcome determined by the shift in the *hox* gene expression. High robustness (low net shift of *hox* genes) may be achieved by a mild (efficient) or a strong (effective) response in the RA metabolic network gene expression (lower quadrants in Fig. 8A). Similarly, the cases of low robustness (high net shift *hox* gene expression) may be accompanied by mild (inefficient) or strong (ineffective) gene expression changes in the RA metabolic network components (upper quadrants in Fig. 9A). In response to increased RA levels, half of the clutches were distributed within the efficient and effective zones (Figs. 9A,B; blue letters). Another five clutches exhibited a substantial shift in *hox* and RA network expression in increased RA, i.e. ineffective response. One clutch was located in an inefficient zone. By contrast, the distribution of clutches in response to RA reduction (DEAB) distributed between the efficient and inefficient quadrants, characterized by limited gene expression shift in the RA metabolic network while yielding a wide range of robustness in *hox* gene expression (Figs. 9A,B; green letters).

We examined if the clutches with similar levels of robustness and overall feedback regulation employ similar strategies for feedback regulation. Interestingly, we found that the identity of the differentially regulated genes in the RA metabolic network can vary between clutches co-localized in the efficiency-efficacy matrix (Figs. 9B,C). For example, in response to increased RA, clutches A, E and F showed similar differential regulation of *dhrs3*, *aldh1a2,* and *cyp26a1*, but diverged in *stra6*, *sdr16c5*, *adhfe1*, *rdh13*, *aldh1a3*, and *cyp26c1* (Figs. 6E and 9C). Similarly, in response to reduced RA, clutches

A, B and D showed downregulation of *dhrs3* and *cyp26a1*, and upregulation of *rbp1*, suggesting feedback aimed at reducing enzymes active in RA reduction, and increasing import of retinol. However, there was high variability in the differential expression of *aldh1a2* and *aldh1a3* to increase RA production (Figs. 6E and 9B,D). Taken together, these results provide strong evidence that the early embryo is capable of robust maintenance of RA levels by mounting a range of network-wide feedback regulatory responses.

Discussion

Robustness of the retinoic acid signaling pathway

Signaling pathway robustness is a central characteristic of all regulatory networks to ensure signaling consistency and reliability overcoming changing environmental conditions or genetic polymorphisms. Gastrula stage *Xenopus laevis* embryos contain RA levels around the 100 nM - 150 nM range (Durston et al., 1989; Kraft and Juchau, 1992; Kraft et al., 1995). Importantly, to strong developmental defects, *Xenopus* embryos, and many other systems, are commonly treated with RA concentrations in the 1 μM to 10 μM range (Sive et al., 1990; Taira et al., 1994). To focus our study of robustness in RA signaling to the physiological range, we show that by increasing the RA levels by as little as 10 nM (about 10%) we could consistently induce expression changes in RA-regulated genes. Interestingly, embryo treatments with physiological RA concentrations exhibit very mild developmental malformations suggesting the induction of compensatory mechanisms to prevent abnormal gene expression (Hollemann et al., 1998; Reijntjes et al., 2005; Sive et al., 1990). We obtained similar results when we reduced the levels of RA by either blocking the biosynthesis (DEAB treatment) or by targeting this ligand for degradation (CYP26A1 overexpression). These observations are supported by multiple loss-of-function studies describing mild

developmental malformations induced by RA signaling reduction (Blumberg et al., 1997; Hollemann et al., 1998; Janesick et al., 2014; Koide et al., 2001; Shabtai et al., 2018; Sharpe and Goldstone, 1997). Therefore, increased or decreased RA levels in the physiological range result in mild developmental defects, suggesting the activation of compensatory mechanisms. We show that one approach to overcome the robustness of RA signaling is to interfere with the metabolic/signaling network at multiple steps (DEAB+CYP26A1) to hamper its ability to efficiently elicit a feedback regulatory response.

We describe a systems biology approach to study the robustness of RA metabolism and signaling based on transient manipulation of ligand levels followed by temporal kinetic, transcriptome-wide analysis of the restoration of normal gene expression patterns. The large transcriptomic differences observed corresponded to normal expression changes resulting from the progression through embryogenesis. Alternatively, the treatments were inefficient, a possibility we could rule out by qPCR screening of all RNA samples for *Hox* expression changes prior to sequencing and subsequent computational analysis of the RNAseq data for individual gene expression shifts. The close clustering of the RA-manipulated (increased or decreased) and control samples in the PCA analysis at each time point further supports the robustness of RA signaling to maintain normal, non-teratogenic, target gene expression levels during early gastrulation.

The RA robustness response emerges from autoregulatory changes in the RA metabolic network

What is the mechanism activated to achieve RA robustness within the physiological range as observed in our study? The early qPCR analysis provided insights on the RA robustness mechanism showing that target genes exhibited

abnormal expression levels at the end of the manipulation (t=0), and some target genes (*hox*) exhibited a speedy return to normal expression after 1.5 hours from the end of the treatment. Genes encoding RA metabolic network components also exhibited abnormal expression levels at t=0, but, their return to normal expression was delayed beyond the time required for the RA targets to reach normal expression. These observations suggest that the RA metabolic network is altered through an RA-dependent feedback regulatory mechanism to restore normal RA signaling and target gene expression. Multiple reports have described the RA-dependent regulation of individual RA metabolic components. Most studies show upregulation of enzymes suppressing or reducing RA signaling like CYP26A1, ADHFe1, and DHRS3, or downregulation of RA producers (anabolic enzymes) like RALDH2 and RDH10 as a result of RA treatment (Chen et al., 2001; Dobbs-McAuliffe et al., 2004; Fujii et al., 1997; Hollemann et al., 1998; Kam et al., 2013; Sandell et al., 2012; Shabtai et al., 2017; Sonneveld et al., 1998; Strate et al., 2009).

To understand the full extent of the kinetic transcriptomic response to transient RA manipulation, we analyzed the recovery kinetics by RNAseq. A discretized pattern analysis of kinetic patterns revealed that only 10 out of the 27 possible patterns were represented in any of the treatment or control samples, and only five patterns were exhibited by a substantial number of genes (n>100). These observations suggest that the RA-regulated genes, direct or indirect, can only exhibit a limited number of possible regulatory responses although intensity differences are possible.

To further understand the regulatory responses of RA targets, we performed a comparative analysis of the discretized patterns using our unbiased approach, COMPACT (Kuttippurathu et al., 2016). This analysis revealed that most genes (75.40%) responding to RA manipulation and recovery exhibited only a response to either increased or decreased RA levels. Only 48 genes out of 553 (8.67%) exhibited

reciprocal responses to increased and decreased RA levels. Interestingly, many of these genes encode RA metabolism or signaling components to possibly maintain non-teratogenic RA levels. The temporal discrepancy in the return to normalcy between targets and network components and the network-wide response taking place at multiple levels provide an integrative view of the feedback regulation and robustness response. Initially, the response to RA changes will most probably be dealt by the actual enzymes and factors already present in the cell. In parallel, the same RA changes will elicit a transcriptional response which for increased RA should involve the upregulation of RA suppressor activities (DHRS3, ADHFe1, and CYP26A1), and the complementary downregulation of RA producing enzymes (RALDH2 and RDH10). This transcriptional response will exhibit a slight delay due to the transcription, translation, and post-translational modifications required.

We observed that some RA network components exhibit oscillatory behavior close to control expression levels, probably as a result of the fine-tuning of the RA signal. The fine-tuning of the RA signal levels coupled to the inherent delay in the transcriptional response could transiently result in the inversion of the overall signaling direction, and oscillatory transcriptional behavior. In a few instances such paradoxical observations of RA signaling outcome following RA manipulation, i.e. overcompensation, have been reported (D'Aniello and Waxman, 2015; D'Aniello et al., 2013; Lee et al., 2012; Rydeen et al., 2015). These observations identify a very dynamic feedback regulatory network continuously fine-tuning itself in response to perturbations.

Asymmetric response to increased and decreased RA levels

Our study provided new insights on the network responses to increased and decreased RA levels. One important conclusion is the observation that these responses

are not symmetrically inverse but rather governed by different regulatory rules. Following a trajectory analysis, we conclude an RA auto-regulatory asymmetry based on:

First, the activation of a response to reduced RA signaling after very slight reduction, while the response to increased RA is only activated above a threshold. Also, the response to reduced RA reaches a lower upper threshold than the response to increased RA. Notably, the network responses show different thresholds depending on the direction of the RA manipulation suggesting a high sensitivity to RA reduction but a lower sensitivity to slight RA increase.

Second, the alignment of the robustness responses, *hox* changes, fitted a diagonal with a negative slope. This negative slope suggested that while a clutch might very efficiently deal with increased RA, i.e. high robustness, the same clutch struggles to compensate for a reduction in RA, i.e. low robustness. The inverse situation, and also more "balanced" clutches were observed suggesting an asymmetric response to RA manipulations.

Third, different clutches with similar robustness levels to increased or reduced RA levels can mount robustness responses by incorporating different components of the RA network. The transcriptomic and HT-qPCR analyses allowed us to perform a detailed determination of the network components comprising a robustness response to uncover its mechanism. The RA metabolic network includes multiple enzymes whose biochemical activity overlaps, e.g., ALDH1A1 (RALDH1), ALDH1A2 (RALDH2), AND ALDH1A3 (RALDH3) enzymes that oxidize retinaldehyde to produce RA (Ghyselinck and Duester, 2019; Kedishvili, 2013; Kedishvili, 2016; Shabtai et al., 2016). The enzymatic efficiencies and expression patterns including location, timing, and intensity may be different across the enzymes (Blentic et al., 2003; Chen et al., 2001; Lupo et al., 2005; Romand et al., 2004; Shabtai et al., 2016). The enzymes

CYP26A1, B1, and C1, or DHRS3, ADHFe1, and RDH13 function to prevent excessive RA signaling (Belyaeva et al., 2008; Belyaeva et al., 2017; Hollemann et al., 1998; Shabtai et al., 2017; Sonneveld et al., 1998). These redundant enzymes and factors can help establish a robustness response using different components.

Finally, genetic polymorphisms probably explain in part the different responses between clutches of the outbred *Xenopus laevis* (Savova et al., 2017). Formally, technical issues could account for some variability. The six additional clutches analyzed by HT-qPCR showed the same variability supporting the involvement of genetic polymorphisms. We mined the Savova et al. (2017) data identifying multiple polymorphisms in RA network components (Supplemental Table S4), supporting the potential contribution of genetic variability in the differential robustness.

Retinoic acid signaling changes due to environmental changes and disease risk

While the RA exposure induces dramatic developmental malformations, treatment of the same embryos with ROL or RAL , the RA precursors, requires higher concentrations (70-100X and 5X respectively) to induce similar defects (Durston et al., 1989; Yelin et al., 2005). Members of the CYP26 family should partially neutralize the RA added or produced (Dobbs-McAuliffe et al., 2004; Hollemann et al., 1998; Sakai et al., 2001). The main difference is that ROL or RAL have to go through RA biosynthesis while RA is already the final active ligand. The oxidation of rol to RAL is a reversible reaction that can reduce substrate availability for the ALDH enzymes, while treatment with RA can only be countered through inactivation by CYP26 enzymes. Therefore, the reduced teratogenicity of the precursors could be the result of reduced conversion to RA as part of the feedback regulation of this network, i.e. robustness, while RA treatment is more restricted in its robustness response.

We obtained evidence of the RA signaling robustness employing several experimental approaches. We observed an efficient response to RA manipulation such that the transcriptomes of treated embryos were not significantly different from controls. All our experiments were performed by either partially inhibiting the endogenous levels of RA, or by slightly increasing (~10% increase) the physiological RA content. Under these conditions, the RA robustness of the embryo efficiently normalizes the transcriptome via regulatory feedback. Based on the comparative analysis of the 12 clutches (genetic backgrounds), we can suggest that robustness efficiency will have a threshold beyond which it will become ineffective in restoring normal RA signaling. Our clutch analysis suggests that this threshold might be strongly dependent on genetic polymorphisms affecting enzymatic activity or gene expression parameters. In support, a threshold or toxicological tipping point for RA signaling was recently described in a cell-based model (Saili et al., 2019). The clutch analysis also showed that the RA network response is also dependent on genetic variability like promoter polymorphisms. Then, the developmental malformations arising from environmental insults on RA signaling largely depend on genetic polymorphisms which will determine the efficiency and threshold of the response and the actual network components comprising such a response.

Materials and Methods

Embryo culture

Xenopus laevis frogs were purchased from Xenopus I or NASCO (Dexter, MI or Fort Atkinson, WI). Experiments were performed after approval and under the supervision of the Institutional Animal Care and Use Committee (IACUC) of the Hebrew University (Ethics approval no. MD-17-15281-3). Embryos were obtained by *in vitro*

fertilization, incubated in 0.1% MBSH and staged according to Nieuwkoop and Faber (1967).

Embryo Treatments

all-trans Retinoic acid (RA), Dimethyl sulfoxide (DMSO), and 4-Diethylaminobenzaldehyde (DEAB), were purchased from Sigma-Aldrich (St. Louis, Missouri). Stock solutions of RA, and DEAB, were prepared in DMSO. Two-hour treatments of 10 nM RA, or 50 μM DEAB, were initiated during late blastula (st. 9.5) and terminated at early gastrula (st. 10.25) by three changes of 0.1% MBSH and further incubation in fresh 0.1% MBSH for the desired time.

Expression analysis

For each sample, 5-10 staged embryos were collected and stored at -80°C. RNA purification was performed using the Bio-Rad Aurum Total RNA Mini Kit (according to the manufacturer's instructions). RNA samples were used for cDNA synthesis using the Bio-Rad iScriptTM Reverse Transcription Supermix for RT-qPCR kit (according to the manufacturer's instructions). Quantitative real-time RT-PCR (qPCR) was performed using the Bio-Rad CFX384 thermal cycler and the iTaq Universal SYBR Green Supermix (Bio-Rad). All samples were processed in triplicate and analyzed as described previously (Livak and Schmittgen, 2001). All experiments were repeated with six different embryo batches. qPCR primers used are listed in (Supplemental Table S2).

RNASeq data analysis

Sequencing was performed at the Thomas Jefferson University Genomics Core using Illumina HiSeq 4000. Reads were mapped to the genome using the Xenopus laevis 9.1 genome with STAR alignment and a modified BLAST and FASTQC in the NGS pipeline (STAR average mapped length: 142.34). Annotation of the mapped sequences using *verse* identified 31535 genes. Raw counts were further filtered for non-zero variance across all samples resulting in 31440 scaffold IDs.

High throughput qPCR

cDNA samples were directly processed for reverse transcriptase reaction using SuperScript VILO Master Mix (Thermo Fisher Scientific, Waltham, MA), followed by real-time PCR for targeted amplification and detection using the Evagreen intercalated dye-based approach to detect the PCR-amplified product. Intron-spanning PCR primers were designed for every assay using Primer3 and BLAST for 24 genes from Retinoic Acid metabolism and target pathway (Supplemental Table S2). The standard BioMark protocol was used to preamplify cDNA samples for 22 cycles using TaqMan PreAmp Master Mix as per the manufacturer's protocol (Applied Biosystems, Foster City, CA, USA). qPCR reactions were performed using 96.96 BioMark Dynamic Arrays (Fluidigm, South San Francisco, CA, USA) enabling quantitative measurement of multiple mRNAs and samples under identical reaction conditions. Each run consisted of 30 amplification cycles (15 s at 95°C, 5 s at 70°C, 60 s at 60°C). Ct values were calculated by the Real-Time PCR Analysis Software (Fluidigm). Samples were run in triplicate for the 24 genes. The primers in the first pre-amplification group selectively bind for the L homoeologues of the genes. In the second pre-amplification group, primers were selected for the S homoeologues only. Third pre-amplification group binds to all 24 genes. In order to remove the technical variability caused due to L and S gene homoeologues, Ct value for each gene in a

sample was selected as the median value of the three pre-amplification runs. Relative gene expression was determined by the ΔΔCt method. *Gapdh* was used as a housekeeping gene for normalization of the data.

Data normalization and annotation

Data analysis on raw genes count was performed using the R statistical analysis tool version 3.6.0 on a 64-bit Windows platform. For the RNA-seq data, the raw genes counts were first converted into log2-transformed values using the "regularized log (rlog)" transformation from the DESeq2 package, which minimizes differences between samples for rows with small counts (Love et al., 2014). The gene expression data was then normalized across samples against the experimental variation and batch effects using COMBAT method in R using a non-parametric adjustment (Johnson et al., 2007). Following batch correction, the gene list was filtered for a minimum expression threshold to remove genes with normalized counts less than 5 across all 72 samples. The expression data for the remaining genes was normalized using quantile normalization**.** RNA-Seq transcript/array probes IDs were transformed to Official Gene Symbol using merged list from 3 sources: the *Xenopus laevis* scaffold-gene mapping, DAVID Bioinformatics Resource 6.8 (Huang et al., 2009) or AnnotationData package "org.Xl.eg.db" maintained by Bioconductor (Carlson, 2017). The original scaffold IDs were retained along with the Official Gene Symbols for cross-reference purposes.

Differential gene expression analysis

The normalized data was analyzed using an Empirical Bayes Statistical (eBayes) model that considered the following two variables and their interactions: (1)

Treatment (Control, RA, DEAB) and (2) Time post treatment-washout (t = 0, 1.5, 3, 4.5h). Differentially expressed genes were identified based on statistically significant effects of Time, Treatment or an interaction between these two factors. P-values were adjusted for multiple testing using topTable from *limma* (Ritchie et al., 2015) package in R ($q \leq 0.05$). The significance-filtered differential gene expression data was used in an established Principal Component Analysis (PCA) approach using the *prcomp* function implemented in R. The samples were annotated based on a combination of treatment and time, yielding 12 distinct sample groups. For each of the selected PC, expression of 100 top positively-loaded and 100 top negatively-loaded genes was visualized using a heat map.

Dynamic pattern analysis and COMPACT analysis

First, for each time point for each of the treatment conditions, the gene expression data for all six clutches (A,B,C,D,E,F) was averaged. Within treatment groups, RA, DEAB, Control, the gene expression data at time points t=1.5, 3, 4.5h was normalized by subtracting the corresponding 't=0h' group. This average differential gene expression data was then discretized to three levels (+1, 0, −1) based on a fold-change threshold [±2 (up, no or down- regulation)]. Within the three treatment groups, this discretization yielded a dynamic response pattern vector for each gene, encoded by one of 27 (3 levels^3 time-points) possible ordered sets. Counts of genes in each treatment group that follow each of the 27 * 27 (=729) possibilities were compared. Functional enrichment analysis was performed for geneset in various dynamic pattern vectors in the Control conditions, using functions *enrichGO* and *simplify* from the R package "clusterProfiler" (Yu et al., 2012) .

COMPACT analysis of RA and DEAB after normalizing to Control

First, for each time point for each of the treatment conditions, the gene expression data for all six clutches (A,B,C,D,E,F) was averaged. Within treatment groups, RA and DEAB, the gene expression data were normalized to Control at each time point by subtracting the expression of the Control group at the corresponding time point. This Control-normalized differential gene expression data for both treatment groups was then discretized to three levels (+1, 0, −1) based on a fold-change threshold [±log2(1.3) (up, no or downregulation)]. This discretization yielded a dynamic response pattern vector for each gene, encoded by one of 81 (3 levels^4 time-points) possible ordered sets. Subsequently, RA and DEAB groups were compared to count the number of genes corresponding to each of the 81 * 81 (=6561) possibilities; to create an 81 × 81 matrix representing the comparative dynamic response pattern counts (COMPACT) (Kuttippurathu et al., 2016). For a given COMPACT matrix of comparative conditions RA(vs. Control) and DEAB (vs. Control), the element at the ith row and jth column of the matrix contains the number of genes that show an 'i'th pattern in DEAB and 'j'th pattern in RA. For a coarse-grained version of the detailed 81x81 COMPACT, pattern-vector counts for each treatment group were further aggregated based on the first time-point, yielding 9 groups of pattern vectors per treatment group. The pair of treatment groups (RA and DEAB) was then compared to count the number of genes corresponding to each of the 9 * 9 (=81) possibilities.

RA network map and visualization

A schematic representation for the position and functioning of the genes involved in the RA biosynthesis, metabolism, translocation, and transcription was formulated

from the literature. For each time point for each treatment, expression value for each gene was mapped to the corresponding label in the schematics using a color scale.

Gene correlation and clustering analysis - Gene expression data was analyzed for both with and without Control-normalization using Weighted Gene Coexpression Network Analysis (WGCNA) (Langfelder and Horvath, 2008) to identify modules of genes with highly correlated differential expression. We used a soft threshold value of 9 (8 for Control normalized data) to identify the initial gene coexpression modules, followed by a dissimilarity threshold of 0.25 to merge the initial modules into the final set of gene coexpression modules. Identified modules were further correlated with the traits (batch, time, treatment).

Robustness score calculation and principal curve-based trajectory analysis

RNA-seq expression data of (clutches: A,B,C,D,E,F) and additional normalized qPCR expression data (clutches: G,H,I,J,K,L) were independently genewise Z-score transformed. Transformed data was then combined to result into 144 samples (12 clutches * 3 treatments * 4 time points). We further selected two subsets of genes: 1) "RA metabolism genes" [aldh1a2.L, aldh1a3.L, crabp2.L, crabp2.S, cyp26a1.L, cyp26a1.S, cyp26c1.L, cyp26c1.S, dhrs3.L, rbp1.L, rdh10.L, rdh10.S, rdh13.L, rdh14.L, sdr16c5.L, stra6.L]. This set represents the feedback regulatory mechanism utilized in response to the RA levels perturbations. 2) "HOX genes" [hoxa1.L, hoxa1.S, hoxa3.S, hoxb1.S, hoxb4.S, hoxd4.L] representing the phenotypic outcome from the treatment. For each gene set, the PCA scores for the combined data was first calculated using R function *prcomp* and then the scores from the first three principal components (PC1, PC2 and PC3) are used to learn a 3-dimensional principal curve using the function *principal_curve* from R package princurve. Two sets of points are specified for the 'start' parameter for this function, which determines the origin and direction of the

curve: (1) The centroid of the 0h-Control samples from all twelve clutches, and (2) centroid of all the remaining 132 samples. For each geneset (RA metabolism genes and HOX genes), for each clutch (A-L), for each treatment (RA or DEAB), a "net absolute expression shift" can be calculated as the sum of distances along the principal curve, of the treatment samples from the corresponding control samples:

$$s.hox_{cl}^{RA} = \sum_{t} \quad abs\left(\frac{\lambda.hox_{cl}^{RA(t)} - \lambda.hox_{cl}^{Control(t)}}{max(\lambda.hox)}\right)$$

where, $\lambda.hox_{cl}^{RA(t)}$ is the arc-distance from the beginning of the curve, for the "*RA treatment samples*" at time point "*t*", for clutch "*cl*" for the principal curve learned for the geneset "*Hox genes*" (Returned as the parameter "*lambda*" from the principal curve function). The expression shift calculation allows ranking and sorting clutches based on 'HOX shift' ($s.hox_{cl}^{treatment}$) and 'RA Network shift' ($s.metabolism_{cl}^{treatment}$). A higher HOX shift for a clutch for a treatment indicates that the clutch is more robust to that particular treatment/perturbation. Furthermore, clutches were mapped onto the conceptual map across the spectrum of "HOX shift" and "RA Network shift" divided into hypothetical quadrants of Effective or Ineffective regulation, or Efficient or Inefficient regulation.

Availability of supporting data/additional files

The raw and normalized datasets for the RNAseq and HT-qPCR data are available online as Gene Expression Omnibus datasets via SuperSeries GSE154408 containing RNAseq data: GSE154399 and HT-qPCR data: GSE154407.

Funding

RV acknowledges financial support for the project from the National Institute of Biomedical Imaging and Bioengineering grant U01 EB023224 and from the Department of Pathology, Anatomy, and Cell Biology, Thomas Jefferson University. AF acknowledges financial support from the Israel Science Foundation (grant 668/17) and the Wolfson Family Chair in Genetics. RV and AF acknowledge the Pilot Funding grant from Thomas Jefferson University and The Hebrew University of Jerusalem collaborative research program.

Abbreviations Used

COMPACT: Comparative Matrix of Pattern Counts

DEAB: 4-diethylaminobenzaldehyde

PCA: Principal Component Analysis

PCR: Polymerase Chain Reaction

RA: retinoic acid

WGCNA: Weighted Gene Coexpression Network Analysis

Author Contributions

R.V. and A.F. conceived and supervised the study and designed the experiments and analysis methodology. L.B-K., M.G., K.K., and A.F. performed embryo experiments and real-time PCR assessment and developed the figures. A.B. performed the initial analysis of the RNAseq data. S.A. conducted the high-throughput PCR validation of RNAseq results. M.P. conducted the analysis of transcriptomics and HT-qPCR data and performed the network and trajectory analyses and developed the figures. M.P., R.V., and A.F. interpreted the results and drafted the manuscript.

Ethics Declarations

The authors declare no competing interests.

Bibliography

Adams, M. K., Belyaeva, O. V., Wu, L. and Kedishvili, N. Y. (2014). The retinaldehyde reductase activity of DHRS3 is reciprocally activated by retinol dehydrogenase 10 to control retinoid homeostasis. *J. Biol. Chem.* **289**, 14868–14880.

Begemann, G., Schilling, T. F., Rauch, G. J., Geisler, R. and Ingham, P. W. (2001). The zebrafish neckless mutation reveals a requirement for raldh2 in mesodermal signals that pattern the hindbrain. *Development* **128**, 3081–3094.

Belyaeva, O. V., Korkina, O. V., Stetsenko, A. V. and Kedishvili, N. Y. (2008). Human retinol dehydrogenase 13 (RDH13) is a mitochondrial short-chain dehydrogenase/reductase with a retinaldehyde reductase activity. *FEBS J.* **275**, 138–147.

Belyaeva, O. V., Adams, M. K., Wu, L. and Kedishvili, N. Y. (2017). The antagonistically bifunctional retinoid oxidoreductase complex is required for maintenance of all-trans-retinoic acid homeostasis. *J. Biol. Chem.* **292**, 5884–5897.

Billings, S. E., Pierzchalski, K., Butler Tjaden, N. E., Pang, X.-Y., Trainor, P. A., Kane, M. A. and Moise, A. R. (2013). The retinaldehyde reductase DHRS3 is essential for preventing the formation of excess retinoic acid during embryonic development. *FASEB J.* **27**, 4877–4889.

Blaner, W. S. (2019). Vitamin A signaling and homeostasis in obesity, diabetes, and metabolic disorders. *Pharmacol. Ther.* **197**, 153–178.

Blaner, W. S., Li, Y., Brun, P.-J., Yuen, J. J., Lee, S.-A. and Clugston, R. D. (2016). Vitamin A absorption, storage and mobilization. *Subcell Biochem* **81**, 95–125.

Blentic, A., Gale, E. and Maden, M. (2003). Retinoic acid signalling centres in the avian embryo identified by sites of expression of synthesising and catabolising enzymes. *Dev. Dyn.* **227**, 114–127.

Blumberg, B., Bolado, J., Moreno, T. A., Kintner, C., Evans, R. M. and Papalopulu, N. (1997). An essential role for retinoid signaling in anteroposterior neural patterning. *Development* **124**, 373–379.

Carlson, M. (2017). Genome wide annotation for Xenopus. *Bioconductor*.

Chen, Y., Pollet, N., Niehrs, C. and Pieler, T. (2001). Increased XRALDH2 activity has a posteriorizing effect on the central nervous system of Xenopus embryos. *Mech. Dev.* **101**, 91–103.

Corcoran, J., So, P. L. and Maden, M. (2002). Absence of retinoids can induce motoneuron disease in the adult rat and a retinoid defect is present in motoneuron disease patients. *J. Cell Sci.* **115**, 4735–4741.

Cui, J., Michaille, J.-J., Jiang, W. and Zile, M. H. (2003). Retinoid receptors and vitamin A deficiency: differential patterns of transcription during early avian development and the rapid induction of RARs by retinoic acid. *Dev. Biol.* **260**, 496–511.

D'Aniello, E. and Waxman, J. S. (2015). Input overload: Contributions of retinoic acid signaling feedback mechanisms to heart development and teratogenesis. *Dev. Dyn.* **244**,

513–523.

D'Aniello, E., Rydeen, A. B., Anderson, J. L., Mandal, A. and Waxman, J. S. (2013). Depletion of retinoic acid receptors initiates a novel positive feedback mechanism that promotes teratogenic increases in retinoic acid. *PLoS Genet.* **9**, e1003689.

Dobbs-McAuliffe, B., Zhao, Q. and Linney, E. (2004). Feedback mechanisms regulate retinoic acid production and degradation in the zebrafish embryo. *Mech. Dev.* **121**, 339–350.

Durston, A. J., Timmermans, J. P., Hage, W. J., Hendriks, H. F., de Vries, N. J., Heideveld, M. and Nieuwkoop, P. D. (1989). Retinoic acid causes an anteroposterior transformation in the developing central nervous system. *Nature* **340**, 140–144.

Eldar, A., Shilo, B.-Z. and Barkai, N. (2004). Elucidating mechanisms underlying robustness of morphogen gradients. *Curr. Opin. Genet. Dev.* **14**, 435–439.

Fainsod, A. and Kot-Leibovich, H. (2018). Xenopus embryos to study fetal alcohol syndrome, a model for environmental teratogenesis. *Biochem. Cell Biol.* **96**, 77–87.

Feng, L., Hernandez, R. E., Waxman, J. S., Yelon, D. and Moens, C. B. (2010). Dhrs3a regulates retinoic acid biosynthesis through a feedback inhibition mechanism. *Dev. Biol.* **338**, 1–14.

Fujii, H., Sato, T., Kaneko, S., Gotoh, O., Fujii-Kuriyama, Y., Osawa, K., Kato, S. and Hamada, H. (1997). Metabolic inactivation of retinoic acid by a novel P450 differentially expressed in developing mouse embryos. *EMBO J.* **16**, 4163–4173.

Ghyselinck, N. B. and Duester, G. (2019). Retinoic acid signaling pathways. *Development* **146**,.

Grandel, H., Lun, K., Rauch, G.-J., Rhinn, M., Piotrowski, T., Houart, C., Sordino, P., Küchler, A. M., Schulte-Merker, S., Geisler, R., et al. (2002). Retinoic acid signalling in the zebrafish embryo is necessary during pre-segmentation stages to pattern the anterior-posterior axis of the CNS and to induce a pectoral fin bud. *Development* **129**, 2851–2865.

Hartomo, T. B., Van Huyen Pham, T., Yamamoto, N., Hirase, S., Hasegawa, D., Kosaka, Y., Matsuo, M., Hayakawa, A., Takeshima, Y., Iijima, K., et al. (2015). Involvement of aldehyde dehydrogenase 1A2 in the regulation of cancer stem cell properties in neuroblastoma. *Int. J. Oncol.* **46**, 1089–1098.

Hollemann, T., Chen, Y., Grunz, H. and Pieler, T. (1998). Regionalized metabolic activity establishes boundaries of retinoic acid signalling. *EMBO J.* **17**, 7361–7372.

Huang, D. W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57.

Janesick, A., Abbey, R., Chung, C., Liu, S., Taketani, M. and Blumberg, B. (2013). ERF and ETV3L are retinoic acid-inducible repressors required for primary

neurogenesis. *Development* **140**, 3095–3106.

Janesick, A., Nguyen, T. T. L., Aisaki, K., Igarashi, K., Kitajima, S., Chandraratna, R. A. S., Kanno, J. and Blumberg, B. (2014). Active repression by RARγ signaling is required for vertebrate axial elongation. *Development* **141**, 2260–2270.

Janesick, A., Wu, S. C. and Blumberg, B. (2015). Retinoic acid signaling and neuronal differentiation. *Cell. Mol. Life Sci.* **72**, 1559–1576.

Johnson, W. E., Li, C. and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127.

Kam, R. K. T., Shi, W., Chan, S. O., Chen, Y., Xu, G., Lau, C. B.-S., Fung, K. P., Chan, W. Y. and Zhao, H. (2013). Dhrs3 protein attenuates retinoic acid signaling and is required for early embryonic patterning. *J. Biol. Chem.* **288**, 31477–31487.

Kedishvili, N. Y. (2013). Enzymology of retinoic acid biosynthesis and degradation. *J. Lipid Res.* **54**, 1744–1760.

Kedishvili, N. Y. (2016). Retinoic acid synthesis and degradation. *Subcell Biochem* **81**, 127–161.

Kessel, M. (1992). Respecification of vertebral identities by retinoic acid. *Development* **115**, 487–501.

Kessel, M. and Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67**, 89–104.

Kim, H., Lapointe, J., Kaygusuz, G., Ong, D. E., Li, C., van de Rijn, M., Brooks, J. D. and Pollack, J. R. (2005). The retinoic acid synthesis gene ALDH1a2 is a candidate tumor suppressor in prostate cancer. *Cancer Res.* **65**, 8118–8124.

Koide, T., Downes, M., Chandraratna, R. A., Blumberg, B. and Umesono, K. (2001). Active repression of RAR signaling is required for head formation. *Genes Dev.* **15**, 2111–2121.

Kot-Leibovich, H. and Fainsod, A. (2009). Ethanol induces embryonic malformations by competing for retinaldehyde dehydrogenase activity during vertebrate gastrulation. *Dis. Model. Mech.* **2**, 295–305.

Kraft, J. C. and Juchau, M. R. (1992). Correlations between conceptal concentrations of all-trans-retinoic acid and dysmorphogenesis after microinjections of all-trans-retinoic acid, 13-cis-retinoic acid, all-trans-retinoyl-beta-glucuronide, or retinol in cultured whole rat embryos. *Drug Metab. Dispos.* **20**, 218–225.

Kraft, J. C., Kimelman, D. and Juchau, M. R. (1995). Xenopus laevis: a model system for the study of embryonic retinoid metabolism. I. Embryonic metabolism of 9-cis- and all-trans-retinals and retinols to their corresponding acid forms. *Drug Metab. Dispos.* **23**, 72–82.

Kuttippurathu, L., Juskeviciute, E., Dippold, R. P., Hoek, J. B. and Vadigepalli, R. (2016). A novel comparative pattern analysis approach identifies chronic alcohol

mediated dysregulation of transcriptomic dynamics during liver regeneration. *BMC Genomics* **17**, 260.

Langfelder, P. and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559.

Lee, L. M. Y., Leung, C.-Y., Tang, W. W. C., Choi, H.-L., Leung, Y.-C., McCaffery, P. J., Wang, C.-C., Woolf, A. S. and Shum, A. S. W. (2012). A paradoxical teratogenic mechanism for retinoic acid. *Proc Natl Acad Sci USA* **109**, 13668–13673.

le Maire, A. and Bourguet, W. (2014). Retinoic acid receptors: structural basis for coregulator interaction and exchange. *Subcell Biochem* **70**, 37–54.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408.

Lohnes, D., Mark, M., Mendelsohn, C., Dollé, P., Decimo, D., LeMeur, M., Dierich, A., Gorry, P. and Chambon, P. (1995). Developmental roles of the retinoic acid receptors. *J. Steroid Biochem. Mol. Biol.* **53**, 475–486.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

Lupo, G., Liu, Y., Qiu, R., Chandraratna, R. A. S., Barsacchi, G., He, R.-Q. and Harris, W. A. (2005). Dorsoventral patterning of the Xenopus eye: a collaboration of Retinoid, Hedgehog and FGF receptor signaling. *Development* **132**, 1737–1748.

Marshall, H., Nonchev, S., Sham, M. H., Muchamore, I., Lumsden, A. and Krumlauf, R. (1992). Retinoic acid alters hindbrain Hox code and induces transformation of rhombomeres 2/3 into a 4/5 identity. *Nature* **360**, 737–741.

Metzler, M. A. and Sandell, L. L. (2016). Enzymatic metabolism of vitamin A in developing vertebrate embryos. *Nutrients* **8**, pii: E812.

Napoli, J. L. (2017). Cellular retinoid binding-proteins, CRBP, CRABP, FABP5: Effects on retinoid metabolism, function and related diseases. *Pharmacol. Ther.* **173**, 19–33.

Niederreither, K., Subbarayan, V., Dollé, P. and Chambon, P. (1999). Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* **21**, 444–448.

Nieuwkoop, P. D. and Faber, J. (1967). *Normal table of Xenopus laevis (Daudin): A systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis.* Amsterdam: North-Holland Publishing Company.

Nijhout, H. F., Best, J. A. and Reed, M. C. (2019). Systems biology of robustness and homeostatic mechanisms. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **11**, e1440.

Paganelli, A., Gnazzo, V., Acosta, H., López, S. L. and Carrasco, A. E. (2010). Glyphosate-based herbicides produce teratogenic effects on vertebrates by impairing

retinoic acid signaling. *Chem. Res. Toxicol.* **23**, 1586–1595.

Pangilinan, F., Molloy, A. M., Mills, J. L., Troendle, J. F., Parle-McDermott, A., Kay, D. M., Browne, M. L., McGrath, E. C., Abaan, H. O., Sutton, M., et al. (2014). Replication and exploratory analysis of 24 candidate risk polymorphisms for neural tube defects. *BMC Med. Genet.* **15**, 102.

Papalopulu, N., Clarke, J. D., Bradley, L., Wilkinson, D., Krumlauf, R. and Holder, N. (1991). Retinoic acid causes abnormal development and segmental patterning of the anterior hindbrain in Xenopus embryos. *Development* **113**, 1145–1158.

Porté, S., Xavier Ruiz, F., Giménez, J., Molist, I., Alvarez, S., Domínguez, M., Alvarez, R., de Lera, A. R., Parés, X. and Farrés, J. (2013). Aldo-keto reductases in retinoid metabolism: search for substrate specificity and inhibitor selectivity. *Chem. Biol. Interact.* **202**, 186–194.

Reijntjes, S., Blentic, A., Gale, E. and Maden, M. (2005). The control of morphogen signalling: regulation of the synthesis and catabolism of retinoic acid in the developing embryo. *Dev. Biol.* **285**, 224–237.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47.

Romand, R., Niederreither, K., Abu-Abed, S., Petkovich, M., Fraulob, V., Hashino, E. and Dollé, P. (2004). Complementary expression patterns of retinoid acid-synthesizing and -metabolizing enzymes in pre-natal mouse inner ear structures. *Gene Expr. Patterns* **4**, 123–133.

Rydeen, A., Voisin, N., D'Aniello, E., Ravisankar, P., Devignes, C.-S. and Waxman, J. S. (2015). Excessive feedback of Cyp26a1 promotes cell non-autonomous loss of retinoic acid signaling. *Dev. Biol.* **405**, 47–55.

Saili, K. S., Antonijevic, T., Zurlinden, T. J., Shah, I., Deisenroth, C. and Knudsen, T. B. (2019). Molecular characterization of a toxicological tipping point during human stem cell differentiation. *Reprod. Toxicol.* **91**, 1–13.

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J. and Hamada, H. (2001). The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. *Genes Dev.* **15**, 213–225.

Sandell, L. L., Lynn, M. L., Inman, K. E., McDowell, W. and Trainor, P. A. (2012). RDH10 oxidation of Vitamin A is a critical control step in synthesis of retinoic acid during mouse embryogenesis. *PLoS ONE* **7**, e30698.

Savova, V., Pearl, E. J., Boke, E., Nag, A., Adzhubei, I., Horb, M. E. and Peshkin, L. (2017). Transcriptomic insights into genetic diversity of protein-coding genes in X. laevis. *Dev. Biol.* **424**, 181–188.

See, A. W.-M., Kaiser, M. E., White, J. C. and Clagett-Dame, M. (2008). A

nutritional model of late embryonic vitamin A deficiency produces defects in organogenesis at a high penetrance and reveals new roles for the vitamin in skeletal development. *Dev. Biol.* **316**, 171–190.

Shabtai, Y. and Fainsod, A. (2018). Competition between ethanol clearance and retinoic acid biosynthesis in the induction of fetal alcohol syndrome. *Biochem. Cell Biol.* **96**, 148–160.

Shabtai, Y., Jubran, H., Nassar, T., Hirschberg, J. and Fainsod, A. (2016). Kinetic characterization and regulation of the human retinaldehyde dehydrogenase 2 enzyme during production of retinoic acid. *Biochem. J.* **473**, 1423–1431.

Shabtai, Y., Shukrun, N. and Fainsod, A. (2017). ADHFe1: a novel enzyme involved in retinoic acid-dependent Hox activation. *Int. J. Dev. Biol.* **61**, 303–310.

Shabtai, Y., Bendelac, L., Jubran, H., Hirschberg, J. and Fainsod, A. (2018). Acetaldehyde inhibits retinoic acid biosynthesis to mediate alcohol teratogenicity. *Sci. Rep.* **8**, 347.

Sharpe, C. R. and Goldstone, K. (1997). Retinoid receptors promote primary neurogenesis in Xenopus. *Development* **124**, 515–523.

Sive, H. L., Draper, B. W., Harland, R. M. and Weintraub, H. (1990). Identification of a retinoic acid-sensitive period during primary axis formation in Xenopus laevis. *Genes Dev.* **4**, 932–942.

Sonneveld, E., van den Brink, C. E., van der Leede, B. M., Schulkes, R. K., Petkovich, M., van der Burg, B. and van der Saag, P. T. (1998). Human retinoic acid (RA) 4-hydroxylase (CYP26) is highly specific for all-trans-RA and can be induced through RA receptors in human breast and colon carcinoma cells. *Cell Growth Differ.* **9**, 629–637.

Strate, I., Min, T. H., Iliev, D. and Pera, E. M. (2009). Retinol dehydrogenase 10 is a feedback regulator of retinoic acid signalling during axis formation and patterning of the central nervous system. *Development* **136**, 461–472.

Taira, M., Otani, H., Jamrich, M. and Dawid, I. B. (1994). Expression of the LIM class homeobox gene Xlim-1 in pronephros and CNS cell lineages of Xenopus embryos is affected by retinoic acid and exogastrulation. *Development* **120**, 1525–1536.

Topletz, A. R., Tripathy, S., Foti, R. S., Shimshoni, J. A., Nelson, W. L. and Isoherranen, N. (2015). Induction of CYP26A1 by metabolites of retinoic acid: evidence that CYP26A1 is an important enzyme in the elimination of active retinoids. *Mol. Pharmacol.* **87**, 430–441.

Urbizu, A., Toma, C., Poca, M. A., Sahuquillo, J., Cuenca-León, E., Cormand, B. and Macaya, A. (2013). Chiari malformation type I: a case-control association study of 58 developmental genes. *PLoS ONE* **8**, e57241.

Yelin, R., Schyr, R. B.-H., Kot, H., Zins, S., Frumkin, A., Pillemer, G. and Fainsod, A. (2005). Ethanol exposure affects gene expression in the embryonic organizer and

reduces retinoic acid levels. *Dev. Biol.* **279**, 193–204.

Yu, G., Wang, L.-G., Han, Y. and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287.

Figure Legends

Figure 1. Phenotypic robustness of the retinoic acid metabolic pathway. Retinoic acid levels were manipulated in *Xenopus laevis* embryos. Inhibition of RALDH activity with DEAB or CYP26A1 overexpression to render retinoic acid inactive were utilized to reduce levels of this signal. **(A)** Schematic diagram of the retinoic acid metabolic pathway and the steps affected. **(B)** Control embryo at st. 27. **(C)** Embryo injected with capped RNA (0.8 ng) encoding the CYP26A1 enzyme. A lineage tracer (ß-galactosidase RNA) was included to ensure a dorsal injection. **(D)** Embryo treated with DEAB (50 μM) from st. 8.5 until st. 27. **(E)** Embryo treated with DEAB and injected with *cyp26a1* mRNA.

Figure 2. Retinoic acid manipulation in the physiological range. Embryos were treated with increasing concentrations of all-*trans* RA from 1 nM to 1 μM. Treatments were initiated at st. 8.5 and RNA samples were collected at early (st. 10.25) and late (st. 12) gastrula. The response of RA metabolic and target genes was studied by qPCR. **(A)** *hoxb1* **(B)** *cyp26a1* **(C)** *dhrs3* **(D)** *raldh2* (*aldh1a2*) **(E)** *raldh3* (*aldh1a3*) **(F)** *rdh10*. Statistical significance (Student's t-test) was calculated compared to the control group. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ns, not significant.

Figure 3. Kinetics of the recovery from RA manipulation. Embryos were transiently treated with either 10 nM RA (A,B), or 50 μM DEAB (C,D). Treatments were initiated during late blastula (st. 9.5), and washed by early gastrula (st. 10.25). RNA samples were collected at different time points during the recovery period. The response of genes, RA target, and RA metabolic enzymes was studied by qPCR. Statistical significance (Student's t-test) was calculated compared to the expression at the end of the treatment (t_0). *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$; ns, not significant.

Figure 4. Kinetic analysis of retinoic acid signal recovery. (A) Schematic description of the experimental design to study the robustness of RA signaling. (B-D) Principal Component Analysis of all six biological replicates for (B) PC1/PC2, (C) PC1/PC8, and (D) PC1/PC10. (E) Heatmap of gene expression of the top-100 positive and top-100 negative loadings corresponding to PC1, PC2, PC8 and PC10. A subset of the genes are highlighted based on their relevance to early developmental processes.

Figure 5. Comparison of the differential gene expression patterns across the two treatment groups compared to controls. (A) Genes were grouped into 27 discretized expression patterns based on upregulation (yellow), downregulation (blue), and no change (grey) above the 2-fold change threshold at each of the three recovery time points compared to the t_0 sample. The ten dynamic patterns that contained at least one gene are shown. The numbers below each graph exemplify the discretized pattern as a numeric vector (+1, 0, -1). The counts next to the patterns indicate the number of genes that show the corresponding expression pattern in each of the two treatment groups and the controls. (B) The Gene Ontology biological processes statistically enriched in the control group are indicated alongside the pattern counts. Details of statistical analysis results are available in Supplemental Table S1. (C) Venn diagrams to compare the overlap between the two treatments and control, illustrated for three differential gene expression patterns. The number of genes that showed similar differential expression patterns in one or more experimental groups are indicated in the corresponding overlapping regions. Black, control; Orange, DEAB; Blue, RA.

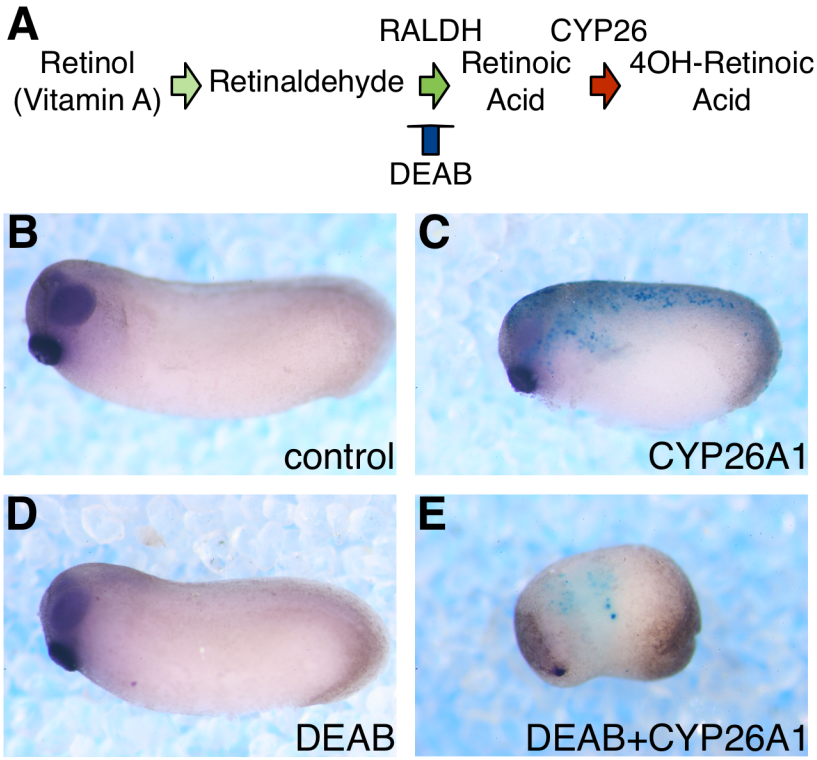
Figure 6. Comparative pattern analysis to uncover genes responding to opposing RA manipulations. (A) Genes were grouped into discretized expression patterns based on upregulation (yellow), downregulation (blue), or no change (grey) compared to the control sample at the same time point. The numbers below the patterns indicate the number of genes that show the corresponding expression pattern in RA vs. control and DEAB vs. control. Only 32 out of 81 theoretically possible (four time points, 3*3*3*3=81) dynamic patterns were exhibited by at least one gene and are included in the figure. (B) COMPACT matrix comparing gene expression changes due to RA and DEAB relative to control. The subset of 32 x 32 patterns with non-zero number of genes in either perturbation group are shown. The gene counts were grouped within related patterns based on the time of initial up- or downregulation. Extended Data 1 contains a version of the COMPACT matrix shown with the gene identifiers corresponding to the counts. (C) Dynamic expression patterns of genes showing opposite changes in response to RA and DEAB treatments (quadrant b). (D) RA metabolic network components and RA target genes. (E) Mapping the differential expression data onto the RA network shown in panel D to highlight the differences in regulation of biosynthesis versus degradation of RA signal between RA and DEAB groups.

Figure 7. Clutch-wise differential expression dynamics of select RA network genes and targets. The clutches are ordered left to right based on the earliest time at which *hoxa1* expression returned to the baseline levels. The data was combined from RNAseq (clutches A-F) and HT-qPCR (clutches G-L).

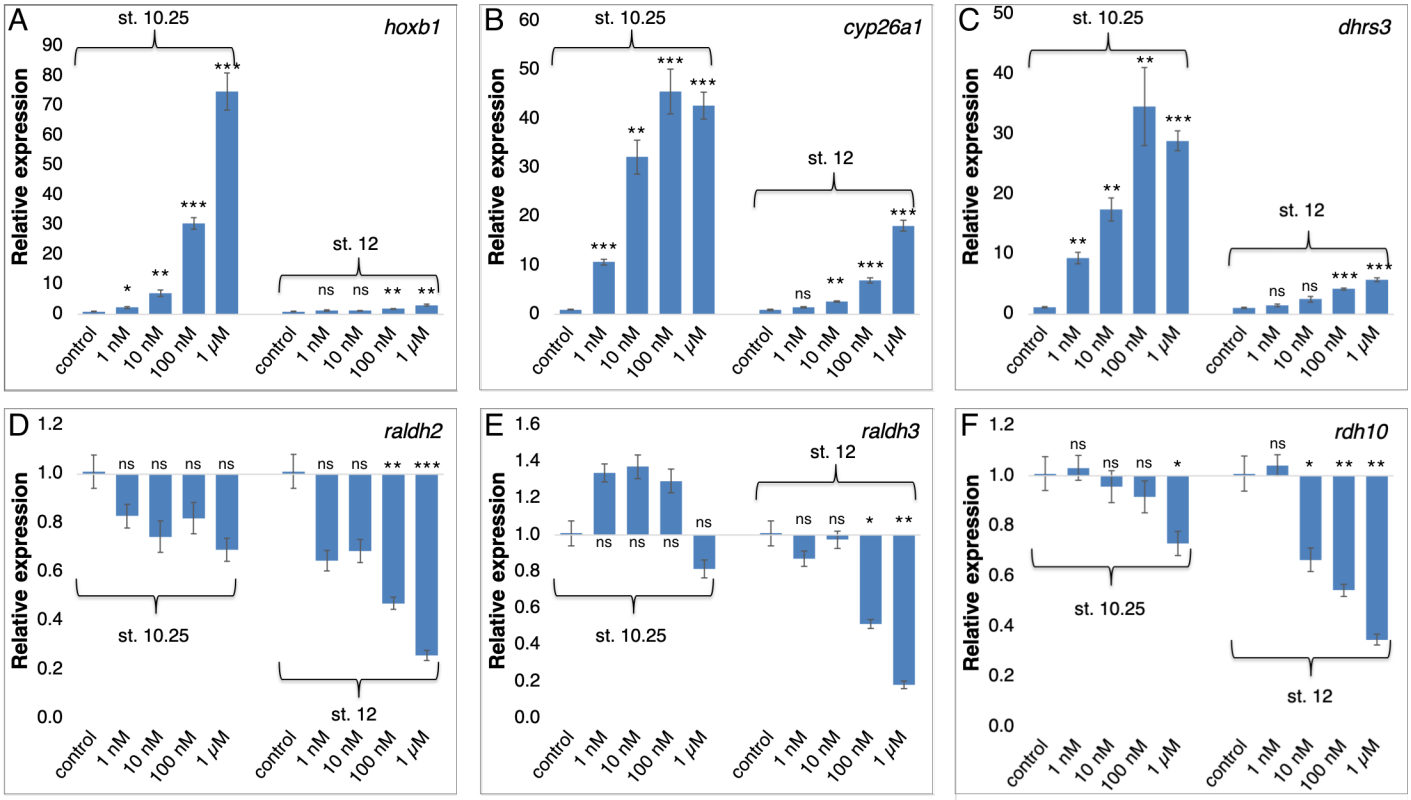
Figure 8. Trajectory analysis to compare the extent of clutch robustness. (A,B) 3-dimensional principal curves for two separate gene sets: RA metabolic genes and *HOX* genes, showing projections of the sample points on the curve for (A) RA and (B) DEAB treatments. Black star indicates the beginning of the curve for the distance measurement along the trajectory. Ranking of clutches is based on the net (absolute) normalized distance of treatment samples from the corresponding Control sample for each time point. Principal curves for clutches C, J, and E. (C,D) Normalized expression shift profile calculated from the principal curve as the arc distance between the treatment and the corresponding control. Clutches are rank-ordered from lowest to highest net expression shift for HOX genes in the RA group. (E,F) Net normalized shift of Hox genes (E) or RA metabolism genes (F) for each clutch over all time points, calculated for both RA and DEAB treatment. Clutches A-F data from RNA-seq , clutches G-L data from HT-qPCR.

Figure 9. Differential robustness to direction of RA change is related to the effectiveness of feedback regulatory action. (A) Schematic diagram of the robustness efficiency matrix. (B) Distribution of clutches across the zones of the robustness efficiency matrix for RA and DEAB treatments relative to the control. The letters indicate the distinct clutches. (C,D) Mapping the differential expression data of clutches A, E, and F onto the RA network shown in Fig. 6D to highlight the heterogeneity of differential regulation of RA network components for clutches closely situated in the robustness efficiency matrix (B).

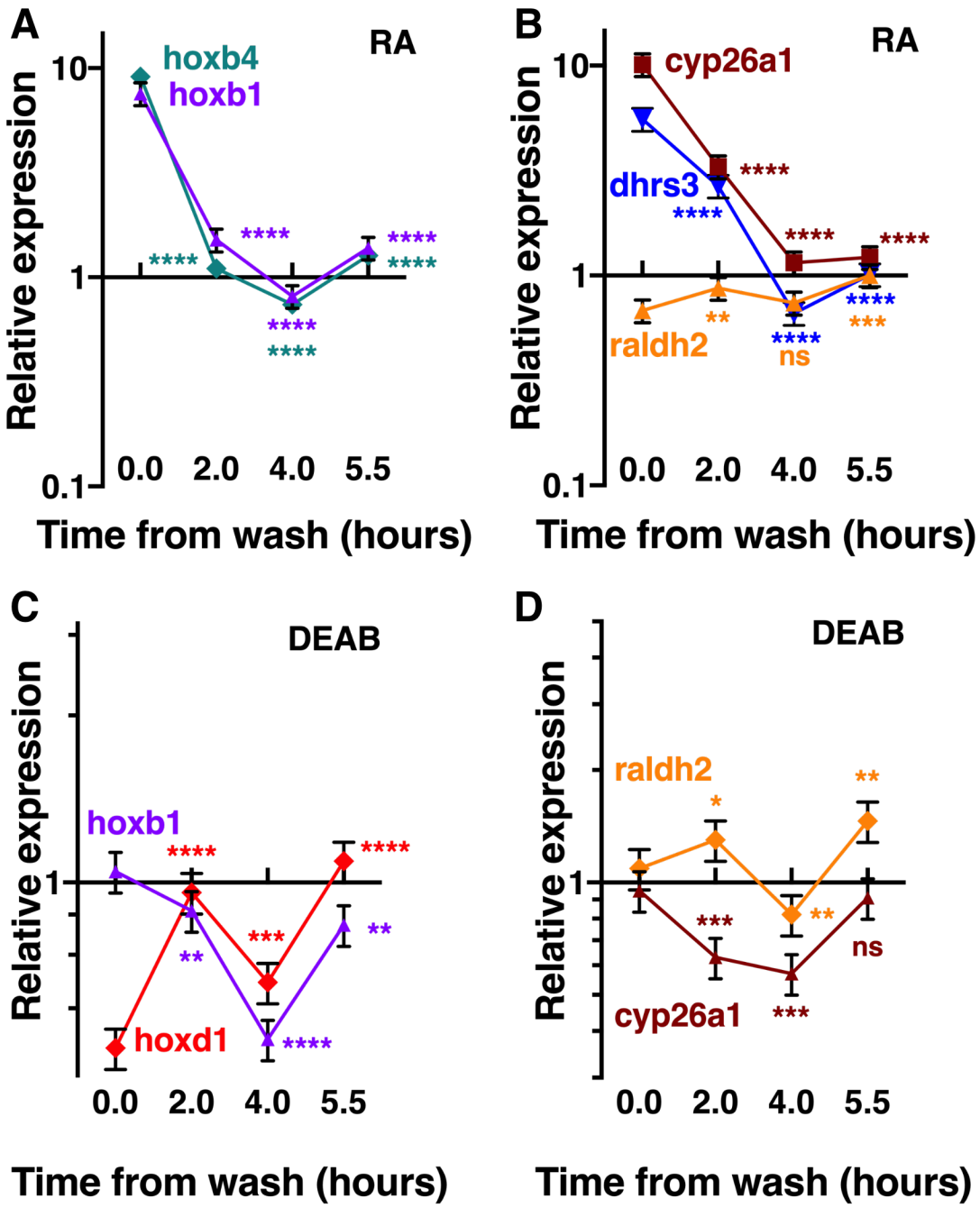
Parihar et al.
Figure 1



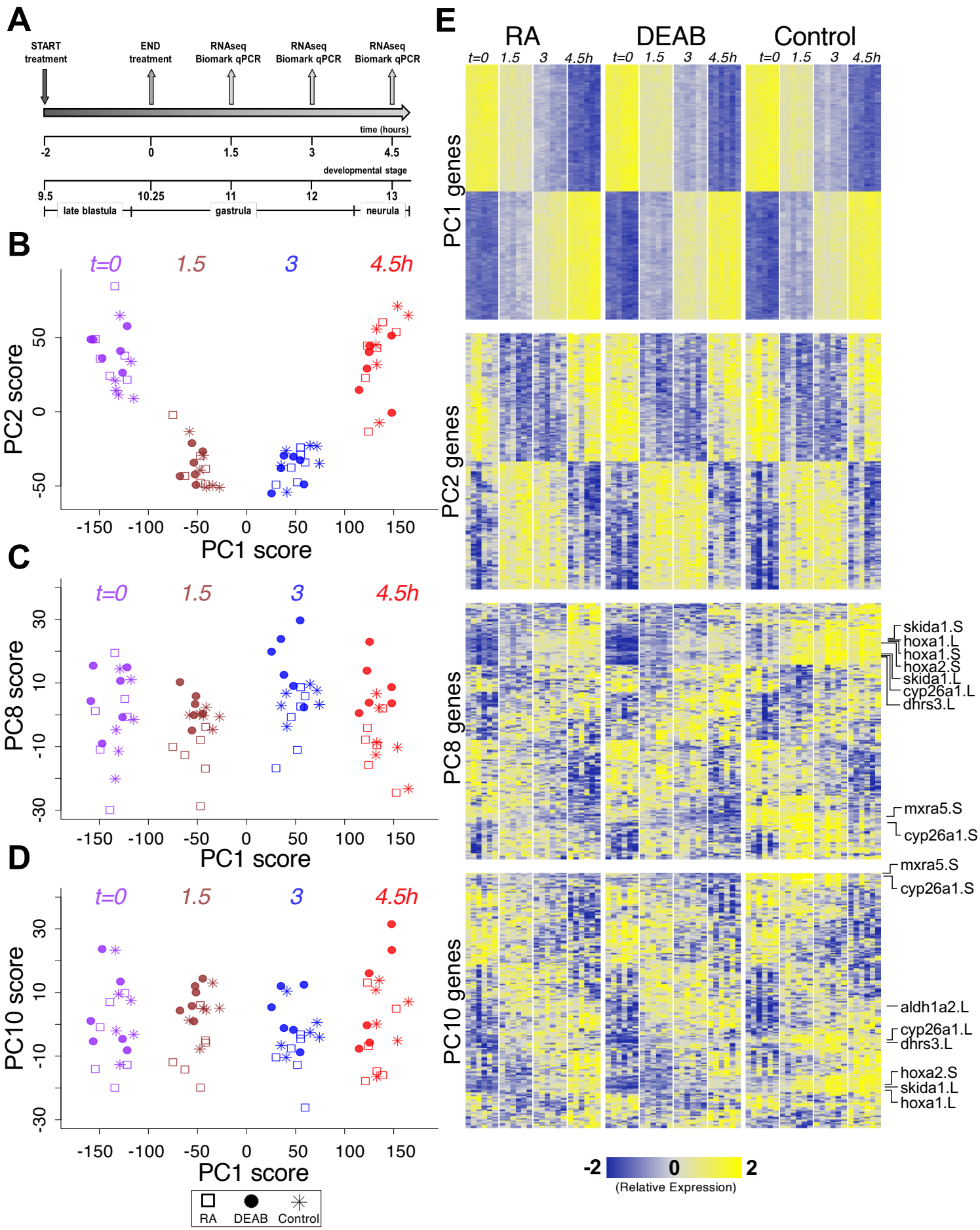
Parihar et al.
Figure 2



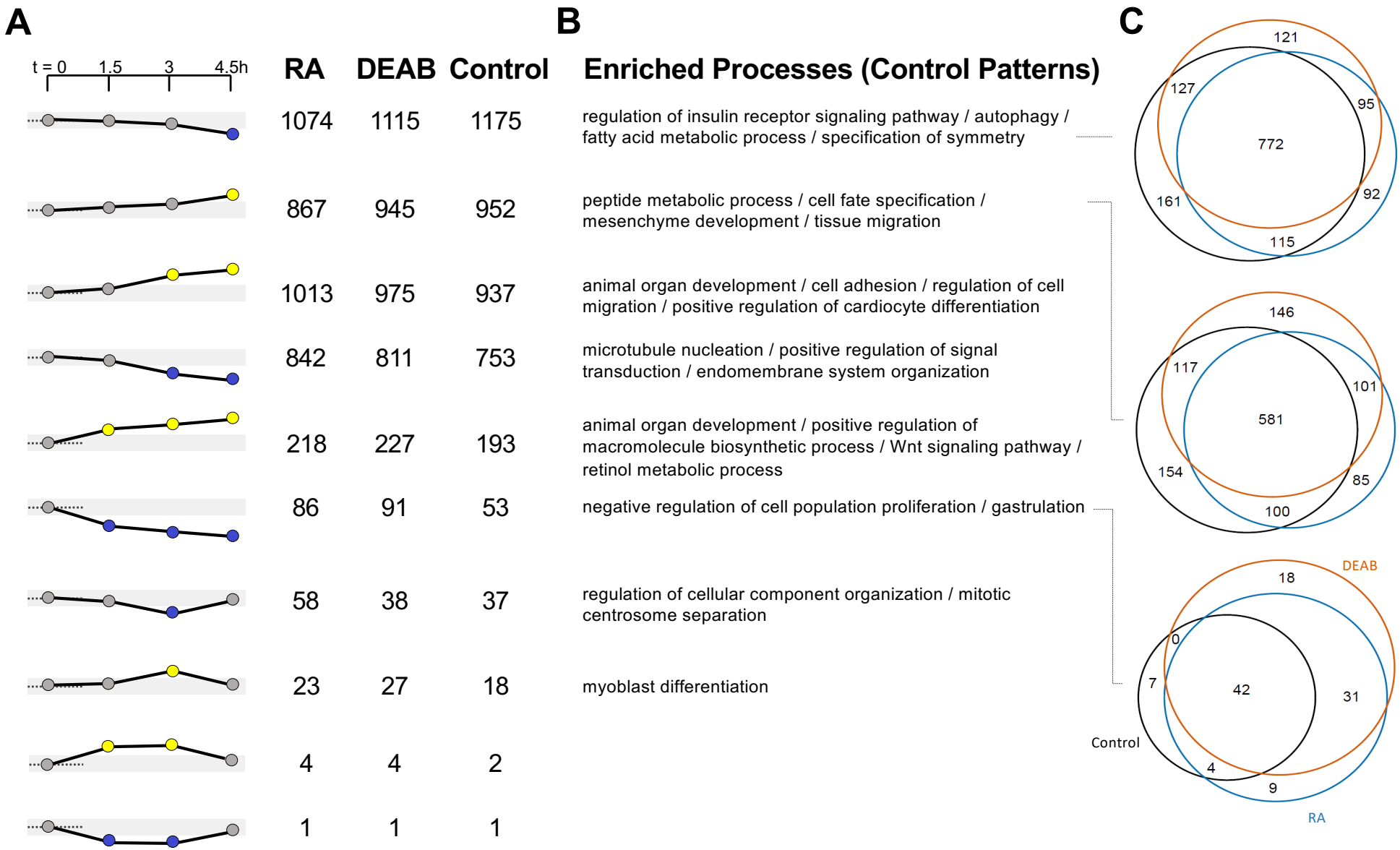
Parihar et al.
Figure 3



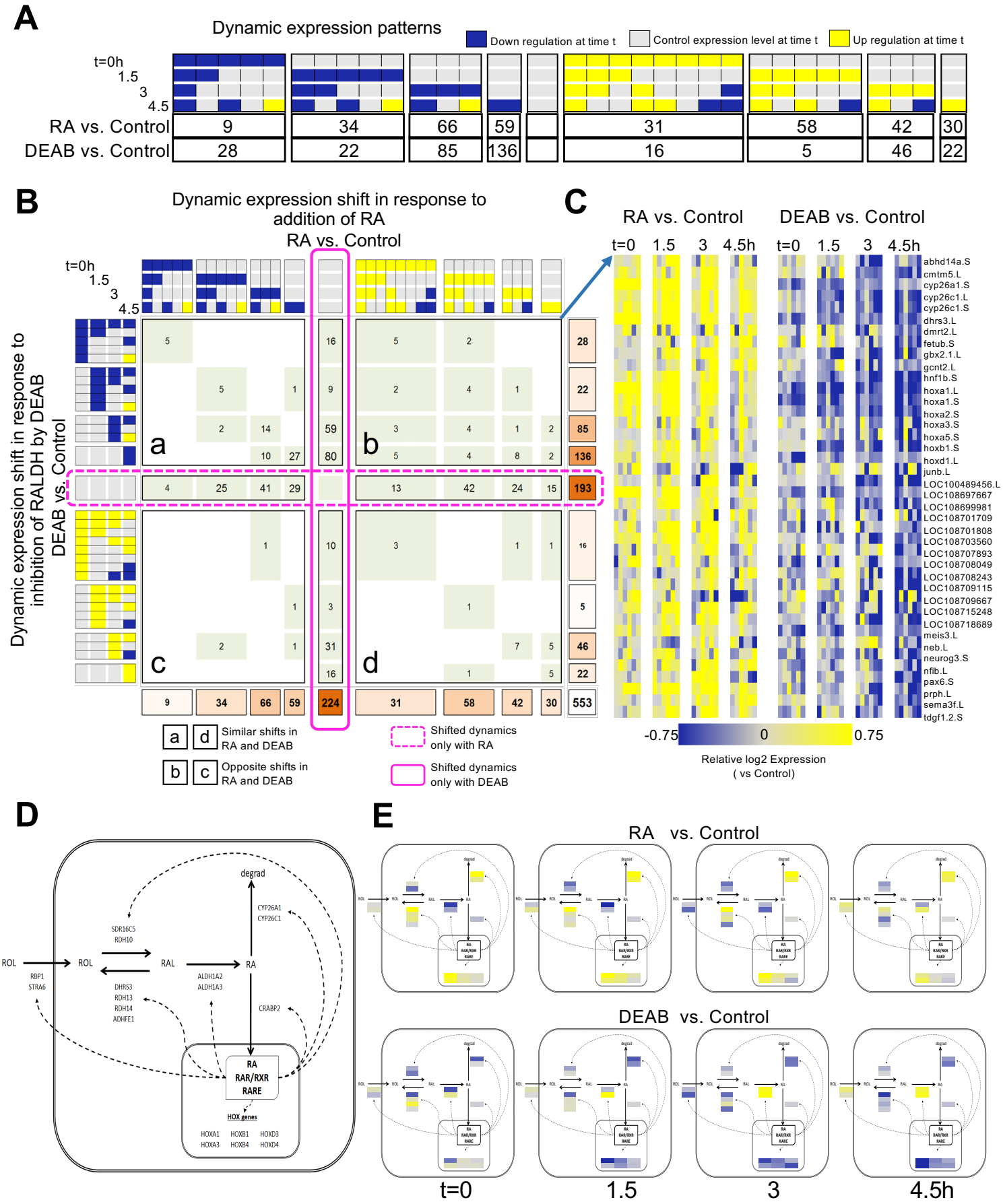
Parihar et al.
Figure 4



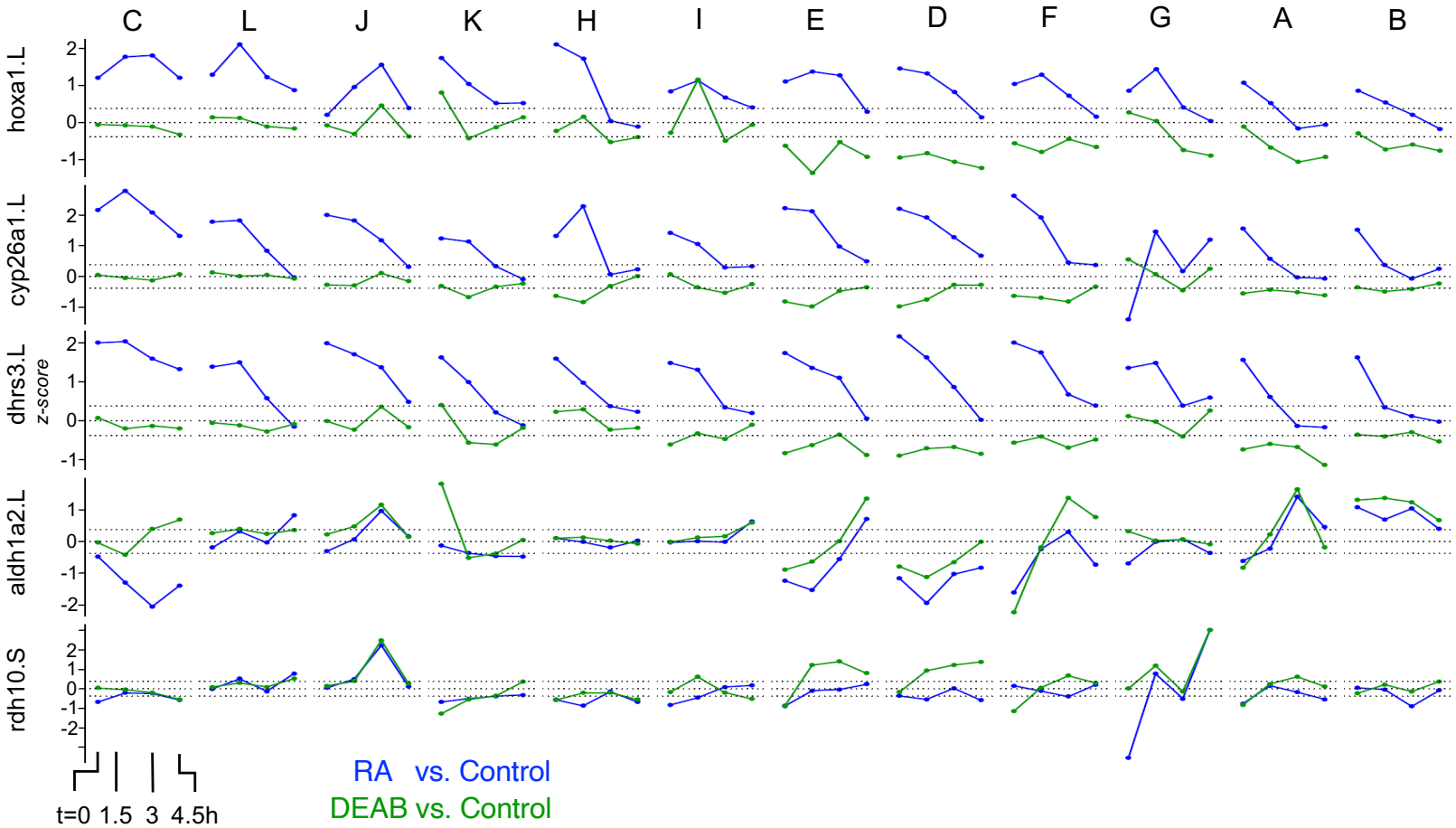
Parihar et al.
Figure 5



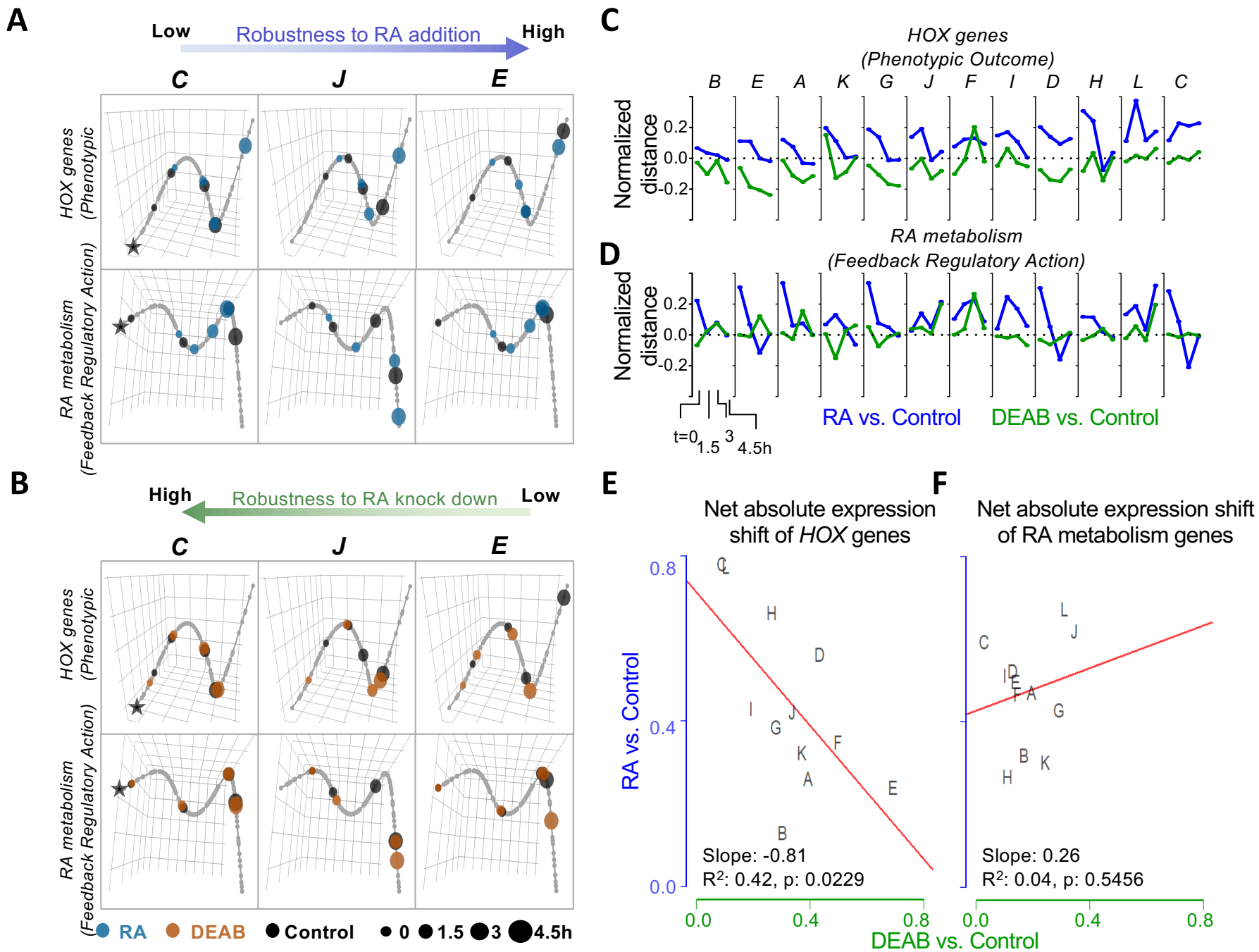
Parihar et al.
Figure 6



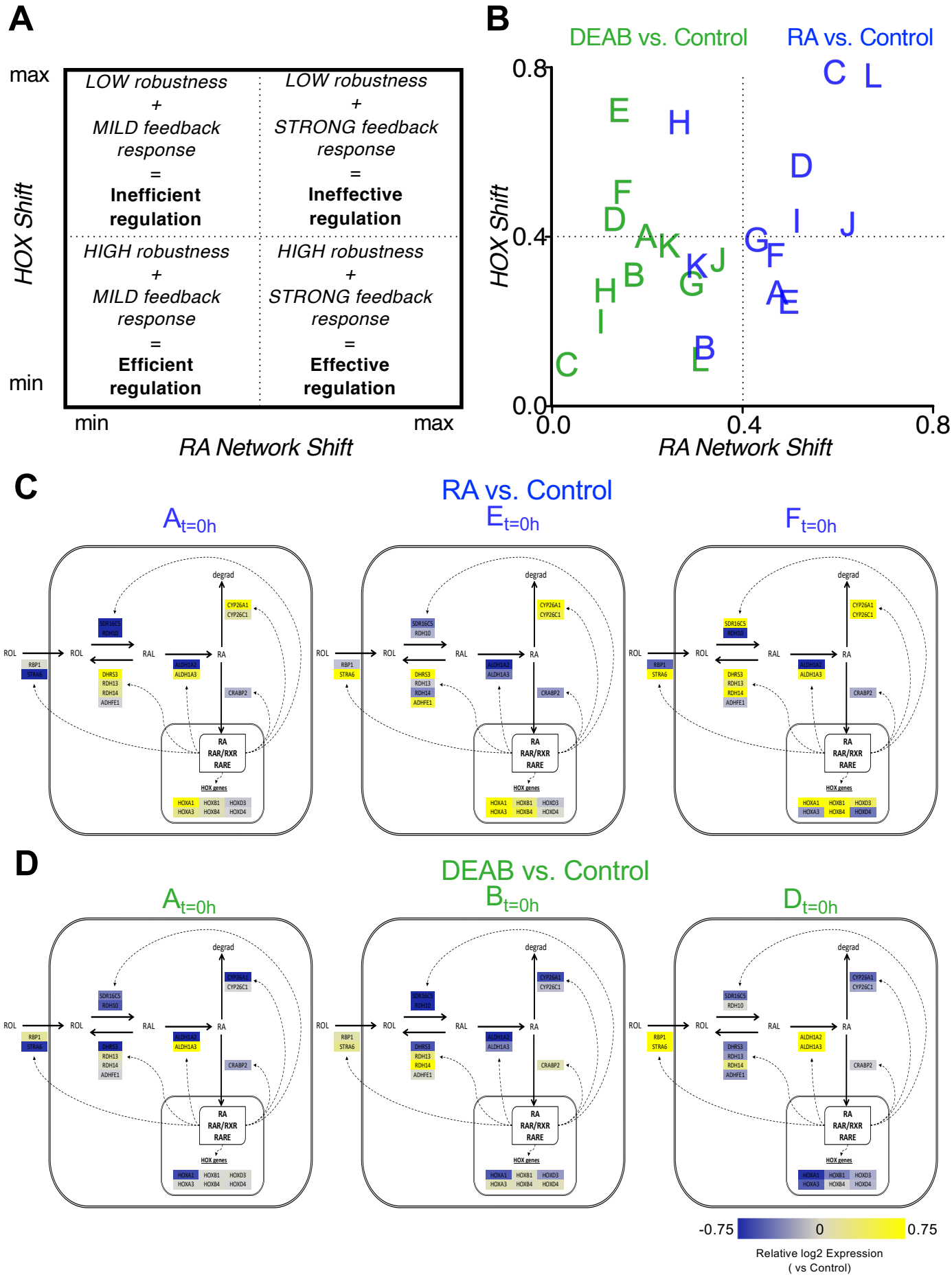
Parihar et al.
Figure 7

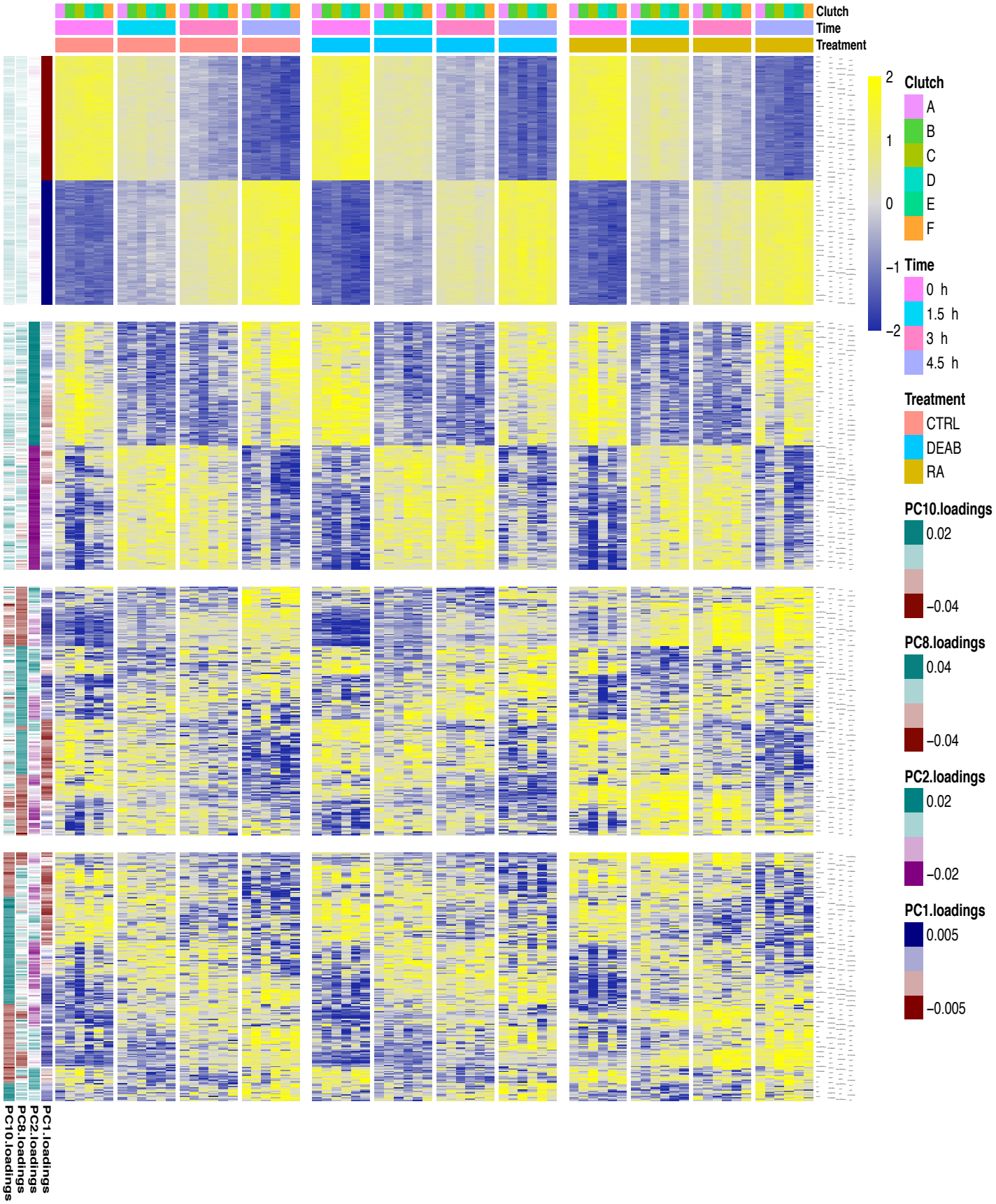


Parihar et al.
Figure 8

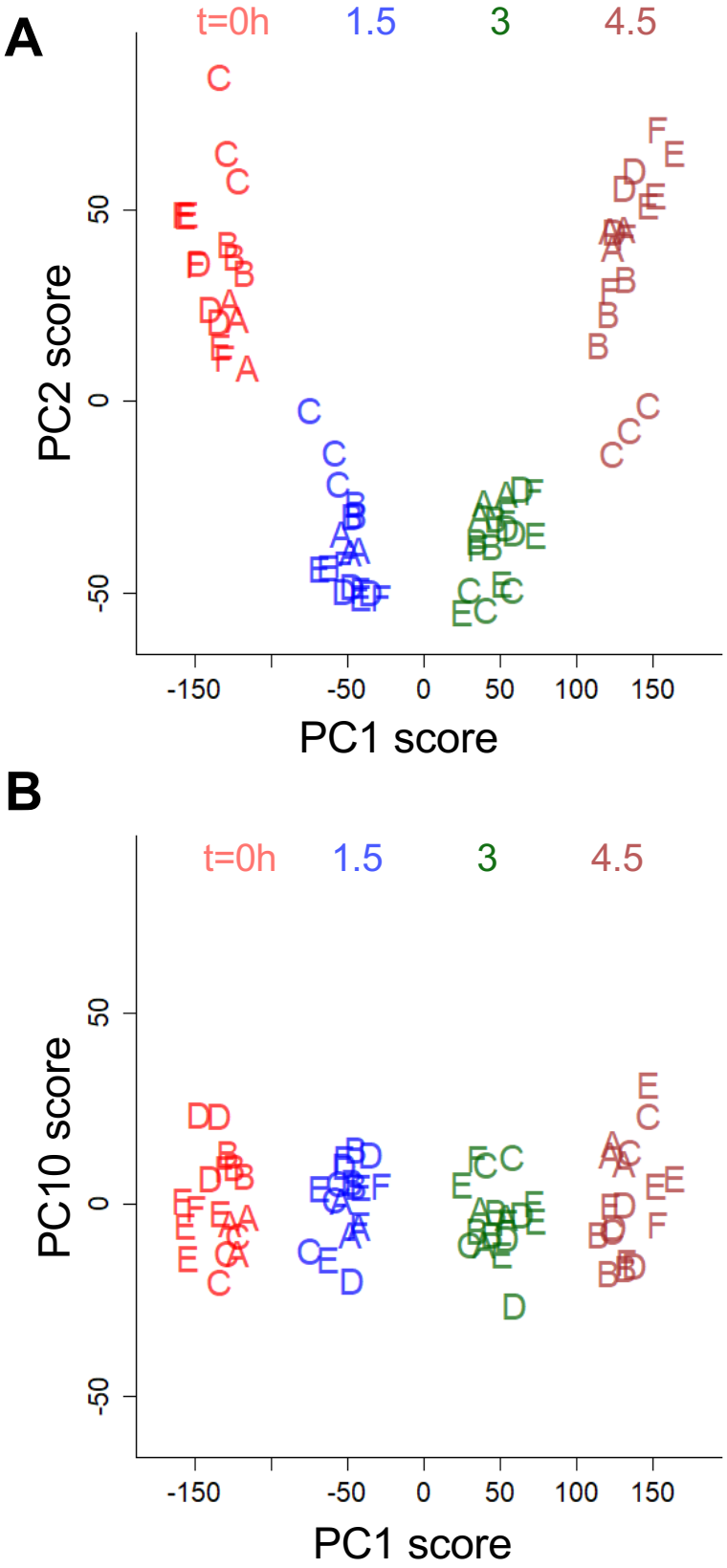


Parihar et al.
Figure 9

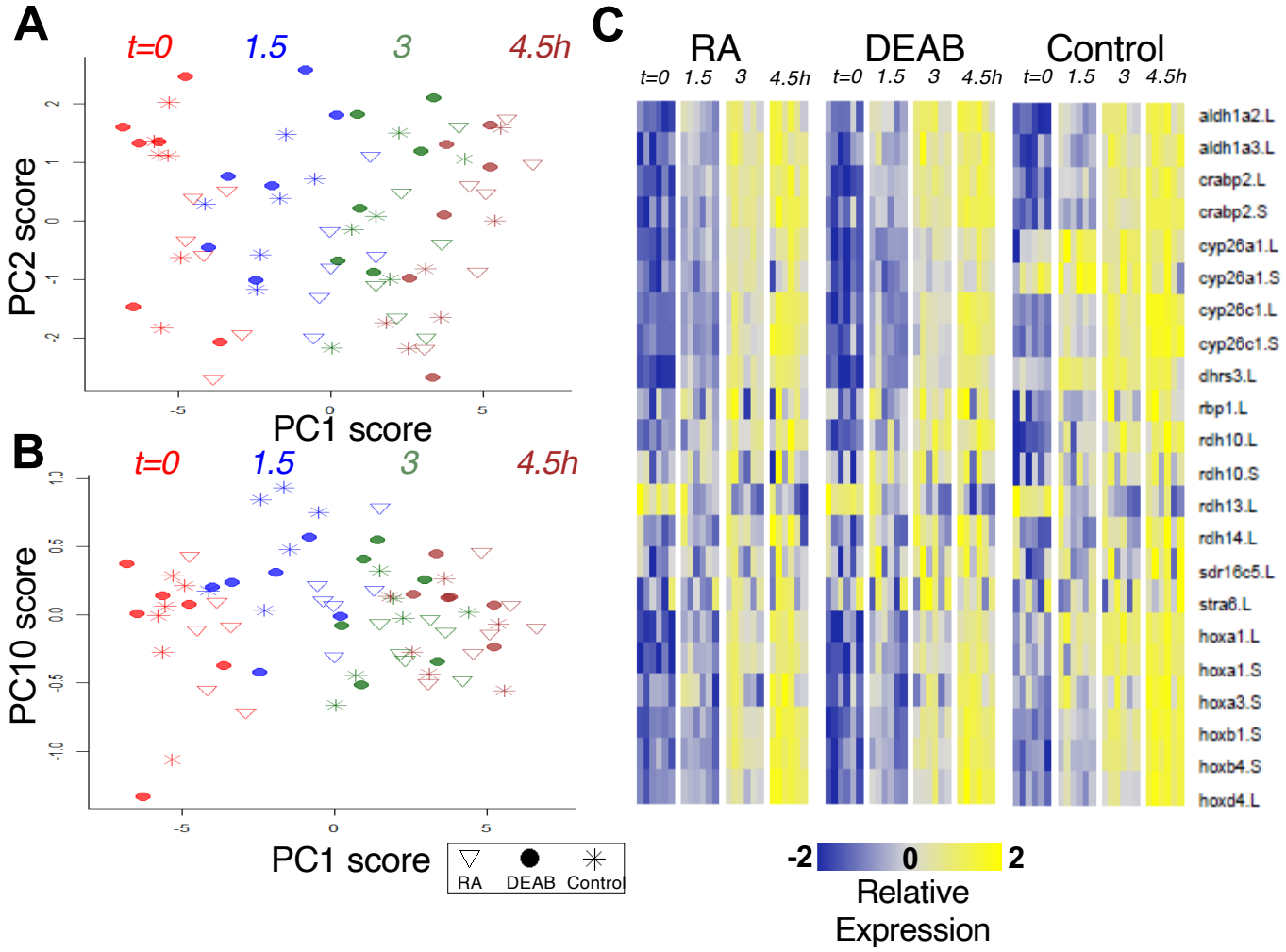




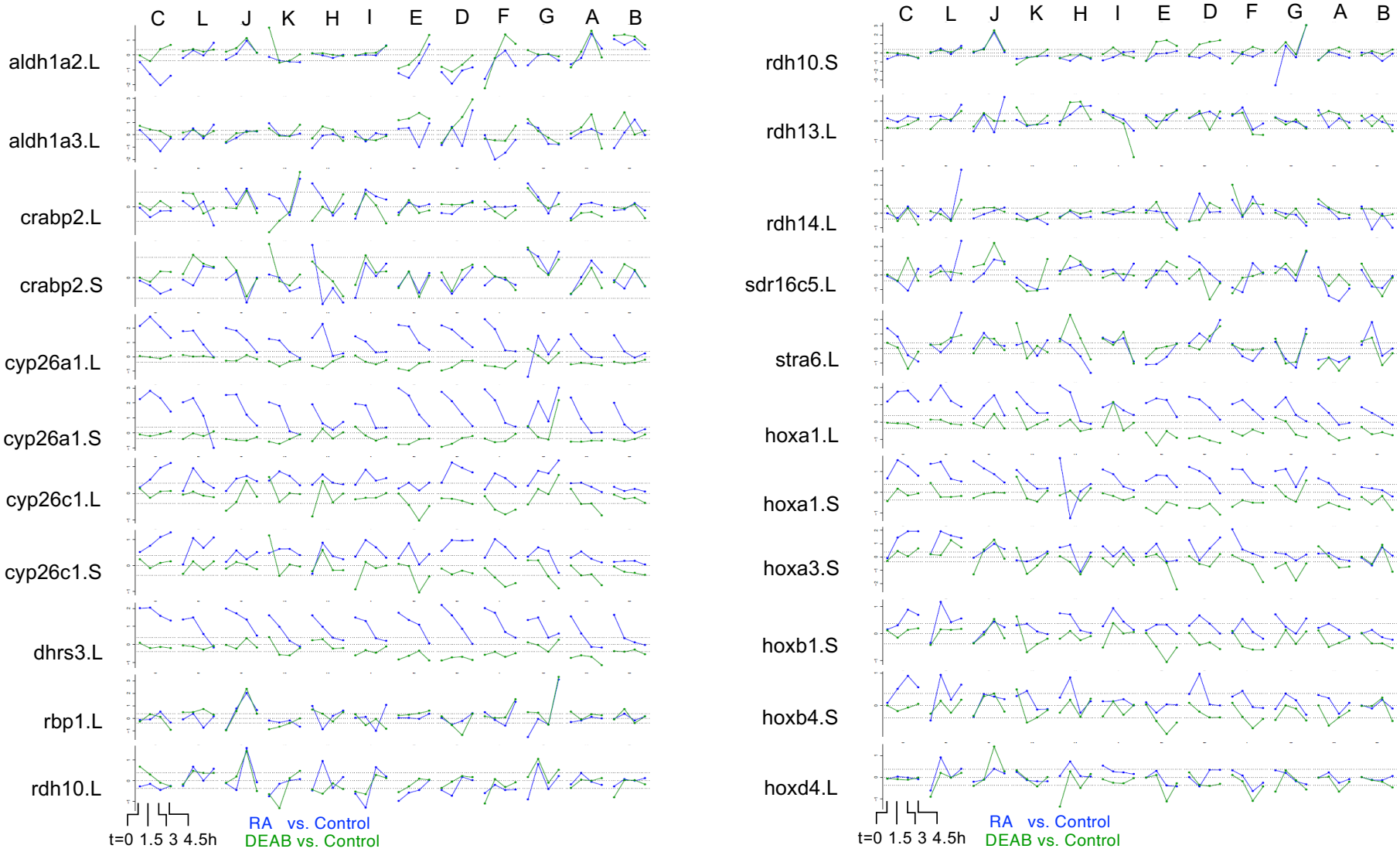
Supplemental Figure S1. Genes affected by RA manipulation. Heatmap of gene expression of the top-100 positive and top-100 negative loadings corresponding to PC1, PC2, PC8 and PC10.



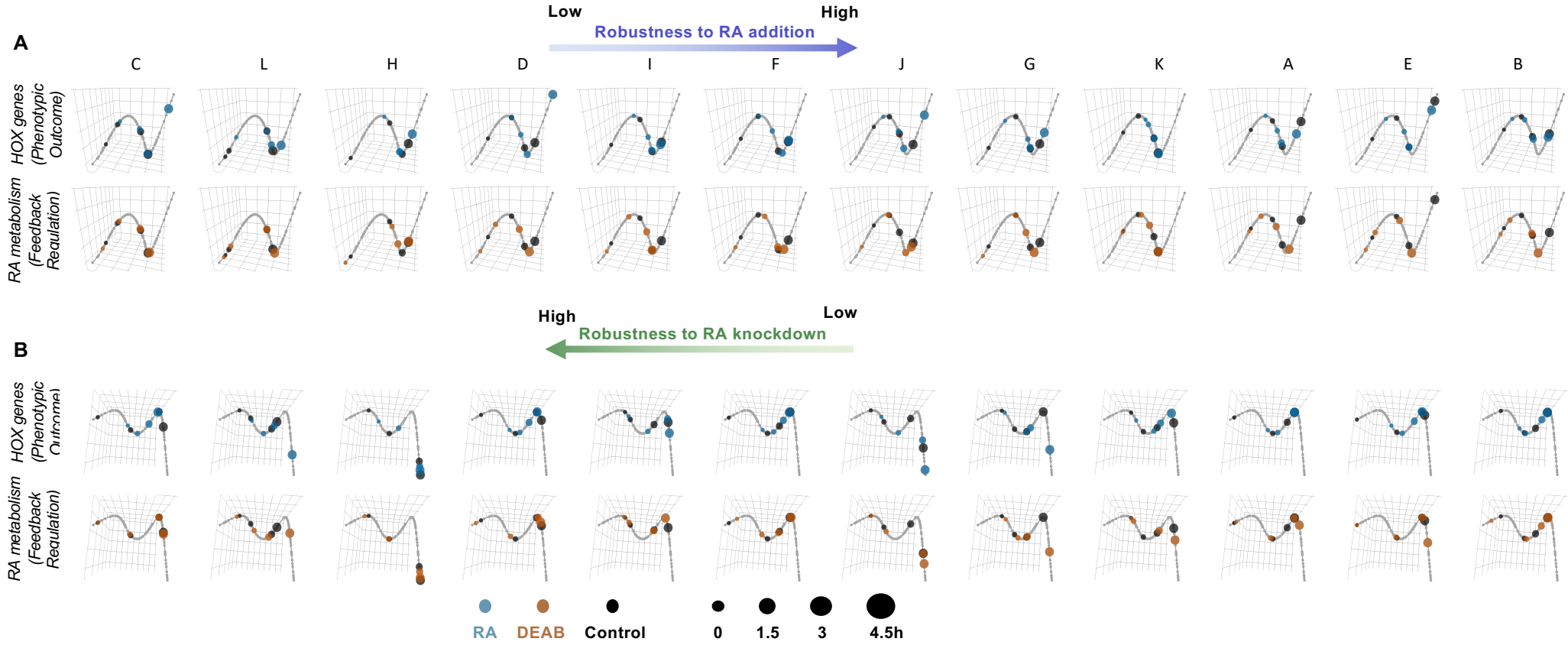
Supplemental Figure S2. Differential robustness to RA manipulation between embryo clutches. Principal Component Analysis of RNAseq data revealed heterogeneity across developmental stages and treatments. (A) Distribution of samples along PC1 and PC2 axes. (B) Distribution of samples along PC1 and PC10 axes. The letters represent the six distinct biological repeats, clutches A-F.



Supplemental Figure S3. HT-qPCR analysis of RA network components. Principal Component Analysis of HT-qPCR data revealed heterogeneity across developmental stages and treatments. (A) Distribution of samples along PC1 and PC2 axes. (B) Distribution of samples along PC1 and PC10 axes. (C) Heatmap of time series differential expression of RA network genes.



Supplemental Figure S4. Clutch-wise heterogeneity in the response of RA network genes following RA manipulation. Clutch-wise differential expression dynamics of RA network genes. The clutches are ordered left to right according to the time taken for recovery of hoxa1.L.



Supplemental Figure S5. Trajectory analysis to compare the extent of multi-gene shift away from and return to control levels and rank clutch-wise robustness. Clutches are ordered left to right according to the extent of net deviation between control and RA treatment along the HOX gene trajectory. High deviation corresponds to low robustness and vice versa. (A) RA and control groups are highlighted along the trajectories. (B) DEAB and control groups are shown.

Supplemental Table S1. List of Gene Ontology annotations with corresponding genes and statistical significance corresponding to Figure 5.

Control Pattern	ID	Description	Total Genes	Pathway Genes	p value	adjusted q value
0 0 0 -1	GO:0046626	regulation of insulin receptor signaling pathway	426	5	2.90E-06	3.99E-03
	GO:0006914	autophagy	426	14	7.65E-05	1.50E-02
	GO:0006631	fatty acid metabolic process	426	11	1.45E-04	2.24E-02
	GO:0006096	glycolytic process	426	7	5.28E-04	3.80E-02
	GO:0008354	germ cell migration	426	3	7.55E-04	4.02E-02
	GO:0009799	specification of symmetry	426	5	1.04E-03	4.74E-02
	GO:0010876	lipid localization	426	13	1.07E-03	4.74E-02
0 0 0 1	GO:0006518	peptide metabolic process	302	44	5.82E-13	2.24E-10
	GO:0001708	cell fate specification	302	4	2.34E-03	2.11E-01
	GO:0060485	mesenchyme development	302	7	3.95E-03	2.68E-01
	GO:0090130	tissue migration	302	3	8.10E-03	4.42E-01
	GO:0010631	epithelial cell migration	302	2	8.83E-03	4.42E-01
0 0 1 1	GO:0048513	animal organ development	320	35	2.34E-08	1.95E-05
	GO:0007155	cell adhesion	320	26	3.21E-07	7.75E-05
	GO:0030334	regulation of cell migration	320	5	6.53E-03	7.35E-02
	GO:1905209	positive regulation of cardiocyte differentiation	320	2	9.88E-03	8.84E-02
0 0 -1 -1	GO:0007020	microtubule nucleation	250	3	3.01E-03	4.81E-01
	GO:0009967	positive regulation of signal transduction	250	9	5.54E-03	4.81E-01
	GO:0098927	vesicle-mediated transport between endosomal compartments	250	2	6.11E-03	4.81E-01
	GO:0010256	endomembrane system organization	250	7	6.13E-03	4.81E-01
0 1 1 1	GO:0048513	animal organ development	83	15	5.49E-07	1.25E-04
	GO:0048732	gland development	83	5	1.42E-06	1.25E-04
	GO:0010557	positive regulation of macromolecule biosynthetic process	83	11	3.29E-06	1.97E-04
	GO:0031328	positive regulation of cellular biosynthetic process	83	11	4.39E-06	1.97E-04
	GO:0016055	Wnt signaling pathway	83	8	2.71E-04	4.57E-03
	GO:0044344	cellular response to fibroblast growth factor stimulus	83	4	4.94E-04	6.21E-03
	GO:0010817	regulation of hormone levels	83	4	1.92E-03	1.84E-02
	GO:0042572	retinol metabolic process	83	2	2.43E-03	2.28E-02
0 -1 -1 -1	GO:0008285	negative regulation of cell population proliferation	15	3	7.57E-05	6.61E-03
	GO:0007369	gastrulation	15	2	5.21E-03	9.93E-02
0 0 -1 0	GO:0007100	mitotic centrosome separation	14	2	1.87E-05	2.16E-03
	GO:0140014	mitotic nuclear division	14	2	5.71E-03	8.00E-02
	GO:0051128	regulation of cellular component organization	14	3	8.47E-03	8.00E-02
	GO:0031338	regulation of vesicle fusion	14	1	8.51E-03	8.00E-02
	GO:0032418	lysosome localization	14	1	9.92E-03	8.00E-02
0 0 1 0	GO:0045663	positive regulation of myoblast differentiation	4	1	2.84E-03	1.71E-02
	GO:0045445	myoblast differentiation	4	1	3.25E-03	1.71E-02

Supplemental Table S2. Weighted Gene Correlation Network Analysis (WGCNA)

Gene ID	Gene Names	Module Numbers	Module Colors	RA-CTRL pattern	DEAB-CTRL pattern
Xelaev18001038	LOC108703560	5	green	1_1_1_1	0_0_-1_-1
Xelaev18002824	sema3f.L	5	green	0_0_1_1	0_0_0_-1
Xelaev18006802	nfib.L	5	green	0_0_1_1	0_0_0_-1
Xelaev18007035	cmtm5.L	5	green	1_1_1_0	0_0_-1_-1
Xelaev18010181	tdgf1.2.S	5	green	0_1_1_0	0_0_0_-1
Xelaev18012533	LOC108708243	5	green	0_0_0_1	0_0_0_-1
Xelaev18013332	prph.L	5	green	1_1_1_1	0_0_0_-1
Xelaev18014991	hnf1b.S	5	green	0_1_1_0	0_0_-1_-1
Xelaev18015206	LOC108709115	5	green	0_0_1_0	0_0_0_-1
Xelaev18024402	pax6.S	5	green	0_0_1_0	0_0_0_-1
Xelaev18024482	LOC108715248	5	green	0_1_1_0	0_0_0_-1
Xelaev18025805	abhd14a.S	5	green	0_1_0_0	0_0_-1_0
Xelaev18030578	LOC108718689	5	green	0_1_1_0	0_0_-1_-1
Xelaev18030982	hoxa1.L	5	green	1_1_1_0	-1_-1_-1_-1
Xelaev18031638	gcnt2.L	5	green	0_0_0_1	0_0_-1_0
Xelaev18033067	hoxa1.S	5	green	1_1_1_0	-1_-1_-1_-1
Xelaev18033068	hoxa2.S	5	green	1_1_0_0	0_-1_-1_-1
Xelaev18033069	hoxa3.S	5	green	1_0_0_0	0_0_0_-1
Xelaev18033072	hoxa5.S	5	green	0_0_1_0	0_0_0_-1
Xelaev18034595	cyp26c1.L	5	green	0_1_1_1	0_-1_-1_-1
Xelaev18035730	dhrs3.L	5	green	1_1_1_0	-1_0_0_-1
Xelaev18036884	cyp26a1.S	5	green	1_1_1_0	-1_-1_0_0
Xelaev18036885	cyp26c1.S	5	green	1_1_1_1	0_0_-1_-1
Xelaev18037155	neurog3.S	5	green	0_1_0_0	0_0_0_-1
Xelaev18037556	LOC108697667	5	green	1_1_1_0	0_-1_0_0
Xelaev18038437		5	green	0_0_1_0	0_0_0_-1
Xelaev18039368_meis3.L	meis3.L	5	green	0_1_0_0	-1_-1_0_0
Xelaev18041724	LOC108699981	5	green	0_1_1_0	0_0_-1_-1
Xelaev18044027	LOC108701808	5	green	0_1_1_0	0_-1_0_0
Xelaev18044028		5	green	1_1_1_0	0_0_0_-1
Xelaev18044734	hoxd1.L	5	green	0_1_0_0	-1_0_0_0
Xelaev18045501	LOC100489456.L	5	green	0_0_1_0	0_0_-1_-1
Xelaev18045983	hoxb1.S	5	green	0_1_0_0	0_-1_-1_-1
Xelaev18047280	gbx2.1.L	5	green	0_1_1_1	0_-1_-1_-1

Supplemental Table S3. List of PCR primers corresponding to the *hox* genes and RA metabolic network for RT-qPCR and HT-qPCR analysis.

Gene	Forward primer	Reverse primer
RT-qPCR		
dhrs3.L	CAGGCGCAAGAAATCCTAAG	CAAAGGCCACGTTACAGGAT
aldh1a2.L	ATGTTTGCCTGGAAGA	GAGAGCAGTGAGCGGA
cyp26a1.S	CGATTCCTCAAGGTTTGGCTTCA	ATTTAGCGGGTAGGTTGTCCACA
hoxb1.L	TTGCCCCAGTGCCAATGAC	TCCCCCTCCAACAACAAACC
hoxd1.S	TTCTTGCGGGGATGTTTTAG	CCGACTGGCATAAAGGAATG
hoxb4.S	CCAAGGATCTGTGCGTCAA	GCAGGATGGAGGCGAACT
gapdh	GCTCCTCTCGCAAAGGTCAT	GGGCCATCCACTGTCTTCTG
HT-qPCR		
aldh1a2.L	ATGTTTGCCTGGAAGATTGC	GAGAGCAGTGAGCGGAGTCT
cyp26c1.L	AAACGGGTTCCTTTCTGTGT	GCTTCGATTTACCCTACACTCTT
hoxb1.S	CCAACTTCACGACCAAACAA	GTGGCTGCGATCTCTACTCTC
hoxd4.L	TTCCCTACCATCATTCCTTTC	GAGTCATTATTTCCTGCTTTCTTT
rbp1.L	TGGAAATGCGAGTAGGAGATG	GGGATGGTGGTTTATTGTGTG
rdh13.L	CAAGTGTCTACCTGGCTGTTG	CCCAGAGTTTCCTTGCAGTT
rdh14.L	TGCCCGTACACAAAGACAGA	GAGACCAAGGAGGTGGTGAG
aldh1a3.L	TAAAGCCCTGTCTGTTTCT	CATACTCTCCAAGTTCCCTT
crabp2.L	AGCCACCCAAAGAAGACATAC	CGATAAGAAACGAAAGCAGAAA
crabp2.S	TCAAAGGAGATGGACCCAAGA	ATCAGCAGTCATGGTCAGGATAAG
cyp26a1.L	TCGAGGTTCGGCTTCATC	CGGCACAATTCCACAACA
cyp26a1.S	CCGCTTTCTAACGCCACTT	CACAATTCCACCACGAACAC
cyp26c1.S	AGCTCTGGTCCTTGAGATGG	AGCCAATGCAGAGTTTCTCC
dhrs3.L	CAGGCGCAAGAAATCCTAAG	CAAAGGCCACGTTACAGGAT
hoxa1.L	CCGCTCACTATATCCACCATTC	TGGCAGGAGAACGACAAAC
hoxa1.S	AATTATGAGATGATGGAATGGTAAA	TGACTGTAAACACCTAGTAAATGAGAG
hoxa3.S	AGCAGGGCAATGGAGTTT	GGAGGGTCCACCAGAATTAG
hoxb4.S	CCAAGGATCTGTGCGTCAA	GCAGGATGGAGGCGAACT
hoxd3.S	CTTCCCAGTCCACAATGAATAAA	GCTCTTCTCCTCGCAGTTCTC
rdh10.L	CGTCTCTTTGCCCTGGAGTTT	CACCATCTCCGCCGTCTC
rdh10.S	TTGCTTGGCCTGTAGAAGAGA	TGCATGGCGAAATAGGAGTAG
sdr16c5.L	TTTGTGGTTCCTTCCCTCTC	GTGCCATCAGTCTCCCTATACC
stra6.L	CCTGCTTACTTTCCTCCTGTTG	GTGGGTGACATTAAAGATTGTAGAGA
gapdh	GCTCCTCTCGCAAAGGTCAT	GGGCCATCCACTGTCTTCTG

Supplemental Table S4. RA network component polymorphisms between *Xenopus* strains[1]

Gene name	*Gene ID*	*aa position*	*aa1*	*aa2*
aldh1a2	18020673m	240	P	S
cyp26c1	18034595m	388	F	I
rbp1	18027898m	26	T	I
rbp1	18027898m	149	E	K
rbp1	18027898m	27	H	R
rdh10	18033799m	133	H	R
rdh13	18026677m	14	C	F
sdr16c5	18032037m	35	A	T
stra6	18018036m	610	S	N

[1]based on (Savova et al., 2017)