

NUCLEAR RECEPTORS

Discovering relationships between nuclear receptor signaling pathways, genes, and tissues in Transcriptomine

Lauren B. Becnel,^{1*†} Scott A. Ochsner,^{2*‡} Yolanda F. Darlington,¹ Apollo McOwiti,¹ Wasula H. Kankanamge,¹ Michael Dehart,¹ Alexey Naumov,^{1‡} Neil J. McKenna^{2§}

We previously developed a web tool, Transcriptomine, to explore expression profiling data sets involving small-molecule or genetic manipulations of nuclear receptor signaling pathways. We describe advances in biocuration, query interface design, and data visualization that enhance the discovery of uncharacterized biology in these pathways using this tool. Transcriptomine currently contains about 45 million data points encompassing more than 2000 experiments in a reference library of nearly 550 data sets retrieved from public archives and systematically curated. To make the underlying data points more accessible to bench biologists, we classified experimental small molecules and gene manipulations into signaling pathways and experimental tissues and cell lines into physiological systems and organs. Incorporation of these mappings into Transcriptomine enables the user to readily evaluate tissue-specific regulation of gene expression by nuclear receptor signaling pathways. Data points from animal and cell model experiments and from clinical data sets elucidate the roles of nuclear receptor pathways in gene expression events accompanying various normal and pathological cellular processes. In addition, data sets targeting non-nuclear receptor signaling pathways highlight transcriptional cross-talk between nuclear receptors and other signaling pathways. We demonstrate with specific examples how data points that exist in isolation in individual data sets validate each other when connected and made accessible to the user in a single interface. In summary, Transcriptomine allows bench biologists to routinely develop research hypotheses, validate experimental data, or model relationships between signaling pathways, genes, and tissues.

INTRODUCTION

Signaling pathways involving members of the nuclear receptor (NR) superfamily of transcription factors, their regulatory small molecules (ligands), and transcriptional coregulators coordinate regulation of gene expression across multiple physiological systems and organs (1, 2). Over the past 15 years, researchers in this field have frequently used transcriptome-scale expression profiling approaches to characterize the biology of NR signaling pathways (3). Although reuse of these data sets has considerable potential value in filling gaps in knowledge in areas of research that the original investigators did not envisage, various factors have restricted such reuse. Rates of public deposition of these data sets are low (4), and many data sets are available only in unpredictable and inconsistent file types. Those that are available in public archives are presented in formats that are intimidating to many bench biologists and are frequently annotated without regard for metadata standards (5). To make these data more routinely accessible to scientists, we previously developed a search tool, Transcriptomine, that aggregated and annotated transcriptomic experiments relevant to NR signaling (6). Here, we describe advances in the biocuration, scale of coverage, and usability of our resource and illustrate its value with reference to a series of Use Cases of evidence gathering, hypothesis generation, and model testing.

¹Nuclear Receptor Signaling Atlas (NURSA) Informatics, Dan L. Duncan Cancer Center Biomedical Informatics Group, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. ²NURSA Informatics, Department of Molecular and Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.

*These authors contributed equally to this work.

†Present address: Clinical Data Interchange Standards Consortium (CDISC), Austin, TX 78701, USA.

‡Present address: Apple Inc., Cupertino, CA 95104, USA.

§Corresponding author. Email: nmckenna@bcm.edu

RESULTS

Transcriptomine allows bench scientists to discover regulatory relationships between signaling pathways, genes, and tissues

The Transcriptomine tool enables data set discovery, reuse, and attribution through two user interfaces (UIs): (i) browsable data set pages, which connect data sets to their associated journal articles and to data set search engines (Fig. 1, Data Set Pages), and (ii) the gene Regulation Report, which allows visualization of biological relationships between genes, signaling pathways, and tissues across the universe of data points (Fig. 1, Regulation Reports). We designed a UI that would enable seamless, bidirectional navigation by the user between the data set pages and Regulation Reports for a gene. The data set pages can be reached from the associated research articles—where publishers support such links—thereby extending the value of the original studies (Fig. 1A). Gene Regulation Reports are accessible through links embedded in gene lists in the data set pages (Fig. 1B) from gene-centric knowledge databases such as Entrez Gene, GeneCards, and the Pharmacogenomics Knowledgebase (PharmGKB) (Fig. 1C) and through queries constructed in the query form itself (<https://goo.gl/oscYup>) (Fig. 1D). Gene Regulation Report data points link directly to a Fold Change Details (FCD) window that provides the experimental details that led to a fold change value (Fig. 1E) and connects it back to its parental data set (Fig. 1F). To complete the cycle of reuse and attribution, the data set page allows one-click citation of the data set in research manuscripts (Fig. 1G), which, in turn, drives discovery of data sets through the reference lists of articles in which they are cited (Fig. 1H).

Classifying transcriptomic data sets by signaling pathway and tissue or cell line enhances their accessibility to bench researchers

The current version of Transcriptomine introduces two newly developed biocuration steps, namely, classifying experiments by signaling

2017 © The Authors,
some rights reserved;
exclusive licensee
American Association
for the Advancement
of Science.

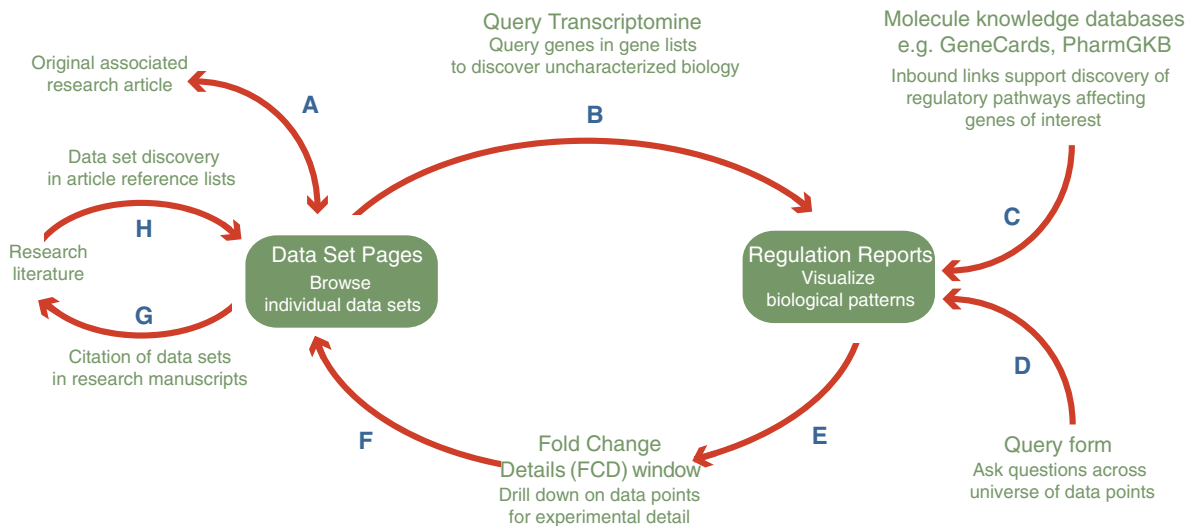


Fig. 1. Transcriptomine enables a cycle of data set discovery, reuse, and attribution to illuminate uncharacterized biology of NR signaling pathways. (A to H) The UI is designed to allow for seamless bidirectional navigation between browsable Data Set Pages and gene Regulation Reports. The Data Set Pages (browsable at the Data Set Directory) link to associated journal articles (A) to extend the value of the original study. The Regulation Report is accessible from the data set pages (B), through links embedded in external gene and small-molecule knowledge databases (C), or directly through a user-configurable query form (D) and enables hypothesis generation through visualization of pathway-gene-tissue relationships. Regulation Report data points link to detailed contextual windows (E) that, in turn, close the loop back to the parental data set (F). Finally, the data set pages enable one-click citation in manuscripts and research proposals (G), which, in turn, drives further discovery of data set citations in article reference lists (H).

pathway and by biosample categories. Transcriptomic data sets in the NR field are based on experiments involving small molecules (physiological ligands, synthetic organics, and others) or genetic manipulations, such as knockout, knockdown, and mutant knockins, of NRs and coregulators. This diversity in experiment design creates obstacles for the user in evaluating evidence for regulation of a gene or gene set by a signaling pathway or in comparing the transcriptomic end points of different signaling pathways. A simple scenario—determining if a gene of interest is regulated by a given signaling pathway—poses two challenges. First, the user must have previous knowledge of all the small molecules or regulatory proteins affecting the pathway, and second, they must run multiple independent queries for all possible experimental manipulations. To eliminate these obstacles to usability, we mapped small molecules and their cognate NRs to a specific pathway (table S1). Tissue specificity is a well-characterized facet of NR signaling pathways (7). Although they perform experiments in specific cell lines or tissues, cell biologists frequently interpret the results in the context of major organs and their associated physiological systems, for example, the role of adipose signaling pathways in metabolism or the role of uterine signaling in female reproduction. To allow users to interact with Transcriptomine using these familiar terms, we therefore assigned biological samples (tissue or cell type) to specific organs (for example, prostate) and their associated mammalian physiological system (for example, male reproduction) (table S2).

The Transcriptomine data set directory enables browsing of data sets and connects them to their associated research articles

The latest version of Transcriptomine contains a reference library of more than 500 data sets, encompassing more than 2000 experimental contrasts and about 50 million data points, with broad coverage of signaling pathways (Fig. 2A) and physiological systems and organs (Fig. 2B). Data sets are browsed in a data set directory (<https://goo.gl/>

8ekO59) that can be filtered using drop-down menus for any combination of signaling pathway, physiological system, and organ.

The data set pages provide for stable digital object identifier (DOI)-based links to external sources, including the associated journal article (Fig. 1A) and citations of data sets in article bibliographies (Fig. 1H) (5). Data set pages (such as <https://goo.gl/wEyHGg>) contain detailed experimental information and are designed with the bench researcher in mind (fig. S1). In addition to the data set name and description (see Materials and Methods), the Overview section (fig. S1A, top) identifies the repository and accession number of the primary data set. The Overview section also allows one-click citation of the data set in a user's reference manager of choice and, to enable the user to readily identify the context in which the data set was originally generated, includes the full citation of the associated research article. Below the Overview section is an Experiments section, partitioned into tabs for (i) Data Points (fig. S1A, bottom), displaying the top induced (red) and repressed (blue) data points in a scatterplot with fold change on the horizontal axis and gene symbols [according to Human Genome Organization (HUGO) gene nomenclature] on the vertical axis, and (ii) Conditions (fig. S1B, top), displaying the experiment name and description and specific parameters such as small-molecule dose or concentration and duration of exposure. A pull-down menu displaying experiment names allows the user to toggle between different experiments of interest (fig. S1A, bottom). Clicking on an individual data point of interest displays a pop-up Fold Change Information (FCI) window containing a link to the Regulation Report for that gene (see section Gene Regulation Reports below and Fig. 1B). This feature allows the user to identify other pathways that might affect a particular gene of interest. The utility of this feature for hypothesis generation and evidence gathering is illustrated in Use Case 3 below. A third section, Related Data Sets, displays Transcriptomine data sets related by regulatory molecule or biosample, options for which can be selected from another pull-down menu (fig. S1B, bottom).

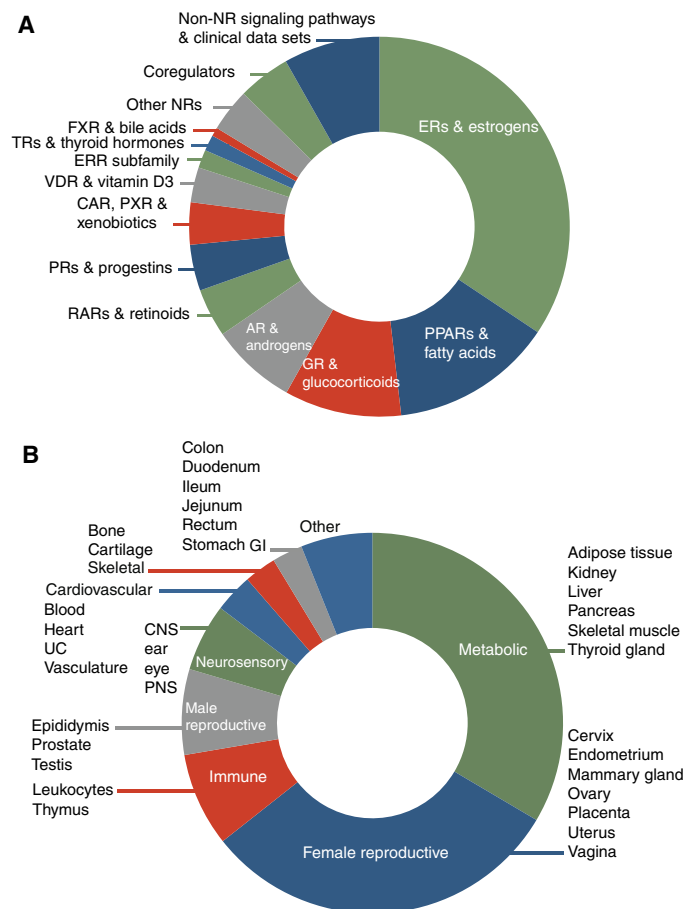


Fig. 2. Breakdown of Transcriptomine data sets by signaling pathway and physiological system and organ. (A) Data sets relevant to ERs and estrogens signaling constitute the largest pathway class in Transcriptomine, followed by, in descending order, studies of signaling by PPARs and fatty acids, GR and glucocorticoids, and AR and androgens. Reflecting cross-talk between NRs and cytoplasmic kinase pathways (141), data sets involving manipulations of cell surface receptors, signaling enzymes, and non-NR transcription factors are an expanding sector of the Transcriptomine data set library. FXR, farnesoid X receptor; TRs, thyroid hormone receptors; VDR, vitamin D3 receptor; CAR, constitutive androstane receptor; PXR, pregnane X receptor; ERR, ER-related receptor; RARs, retinoic acid receptors. (B) Data sets involving female reproductive and metabolic tissue model systems constitute nearly two-thirds of the database. The prominence of female reproductive biosamples reflects the popularity of mammary epithelial cell line models, whereas the large number of metabolic biosamples is due, in part, to our curation of data sets emerging from a TG-GATEs (142), a large-scale toxicotranscriptomic screen in liver and kidney model systems. GI, gastrointestinal; CNS, central nervous system; PNS, peripheral nervous system; UC, umbilical cord.

The query form allows bench scientists to ask biological questions in Transcriptomine

Rather than looking at a specific data set, users often have more open-ended questions in mind. What signaling pathways regulate my gene or function of interest? What genes are most frequently regulated by my pathway of interest in different tissues? How does regulation of my cellular function of interest by a given signaling pathway differ between different tissues? Users can ask these types of questions using the Transcriptomine query form (<https://goo.gl/oscYup>) (fig. S2). The pathway and biosample biocuration described above allowed us to simplify the query form to three principal elements: pathway, biosample category, and

gene(s) of interest (fig. S2). Drop-down menus representing the pathway (fig. S2A) and biosample (fig. S2B) classes serve the dual purposes of (i) conveying the overall scope of the resource and (ii) eliminating possible user confusion or input error associated with manual text input. Selecting a given NR pathway option (fig. S2A) will retrieve all experiments that map to that pathway, whereas selecting a biosample physiological system displays an additional menu that allows the user to select individual organs (fig. S2B). Options for Gene of Interest (fig. S2C) are as follows: Any, Single Gene (text autocomplete allows for disambiguation of synonyms to a single approved gene symbol), Gene List (upload of up to 5000 approved gene symbols), Gene Ontology Term [querying by a user-specified Gene Ontology (8) term], and Disease Term [querying by a disease curated by the Online Mendelian Inheritance in Man (OMIM) curation initiative (9)]. A fourth module (fig. S2D) enables modification of the default P value cutoff (≤ 0.05) for fold change results.

Gene Regulation Reports allow for varied perspectives on pathway-gene-tissue relationships

In the following section, Transcriptomine data points are cited using paired references in which the Transcriptomine data set is immediately followed by its associated research article. Transcriptomine query results are displayed in a gene Regulation Report (figs. S3 and S4), a scatterplot similar to that used in the data set pages. The vertical axis enables four views of the data points, of which Pathway is the default for single gene queries. In the Pathway view for the Transcriptomine *GREB1* Regulation Report (<https://goo.gl/dCHoA6>) (fig. S3 and data file S1, Query 1), top-level Pathways are organized into individual rows identified by a symbol representing the corresponding experiment regulatory molecule(s) (see table S1 and Materials and Methods), making biological patterns in the data readily visible. For example, numerous data points confirmed the published regulation of *GREB1* by the estrogen receptors (ERs) and estrogens pathway (10), the androgen receptor (AR) and androgens pathway (11), and the glucocorticoid receptor (GR) and glucocorticoids pathway (12). In addition, the search results indicated previously uncharacterized regulation of *GREB1* by the peroxisome proliferator-activated receptors (PPARs) and fatty acid-regulated signaling pathway (fig. S3).

Segregation of data points by their corresponding regulatory molecules within a given pathway allows for the comparison of the pharmacology of different small molecules with respect to a given gene in a specific biosample. *GREB1*, for example, was shown by Transcriptomine to be induced in mammary gland model systems by the physiological ER subfamily agonist 17 β -estradiol (17 β E2) (13, 14) and the mammary estrogens genistein (15, 16) and PCB54 but was repressed by the mammary anti-estrogens 4-hydroxytamoxifen and raloxifene (17, 18), as well as by the pure anti-estrogen fulvestrant (19, 20). Similarly, the inclusion of data points from knockout studies provides information on the possible receptor dependence of a given pathway-gene regulatory link. Illustrating this, repression (blue) and induction (red) of *GREB1* in MCF-7 cells by depletion (21, 22) and overexpression (23, 24), respectively, of ER α /ESR1 indicated that this receptor is required for estrogenic regulation of *GREB1* in these cells (fig. S3). Note that according to our previously proposed convention (25), to make our resource accessible to users in both the NR and non-NR fields, individual regulatory molecules (receptors and transcriptional coregulators) are referred to using the symbol in common use in the field and the species-appropriate approved symbol (for example, ER α /ESR1).

Coregulators are a diverse class of molecules that are required by NRs for efficient regulation of gene expression (1, 26). Many coregulators have broad specificity for multiple NRs, so to avoid unduly implying

a specific association with a pathway in regulation of a gene, data points involving manipulations of individual coregulators are categorized in a separate Coregulators section (fig. S3). This segregation allows visualization of potential functional roles of coregulators in regulation of a given gene or group of genes by a specific NR pathway. For example, repression of *GREB1* in HepG2 cells overexpressing a dominant-negative *PGC1 α /PPARGC1A* (27, 28) indicated that induction of *GREB1* by the PPARs and fatty acids signaling pathway involves the transcriptional coregulator *PGC1 α /PPARGC1A* (fig. S3 and data file S1, Query 1).

In addition to the Pathway view, the Regulation Report also allows viewing of data points categorized by biosample, gene, or species. In the Biosample view, individual rows containing all data points from experiments in a specific tissue or cell line are organized into organs and physiological systems (table S2). The Biosample view for the Transcriptome *CA12* Regulation Report (<https://goo.gl/PvmNsD>) (fig. S4) illustrates how this view helps the user identify tissue-specific regulation of gene expression by a pathway. Consistent with its function in bone (29), *CA12* is shown in Transcriptomine to be frequently regulated in bone model systems, principally U2OS cells. In contrast, the abundant data points in mammary gland model systems visualized in the *CA12* Regulation Report indicate a function in this organ that is less well characterized. In the Gene view, alphabetical listing by symbol is the default display for multigene queries. To provide a less visually cluttered UI, we used the HUGO human gene symbol for all orthologs of a receptor. Finally, the Species view enables comparison of patterns of regulation of gene(s) by a pathway between different species.

Transcriptomic approaches are frequently used to evaluate the effect of small-molecule or genetic manipulations on gene expression programs involved in animal or cell models, such as the Zucker diabetic rat or adipogenesis in the NIH 3T3-L1 cells (table S3). When collated in a single interface, control experiments from these data sets are a convenient reference for identifying transcripts that are consistently regulated in these models. To enable such discovery, we categorized these data points in a distinct section of the Regulation Report (fig. S3, Animal and Cell Models). Use Case 2 (see below) illustrates the utility of these model data points in modeling NR pathway–biological process relationships. In addition to these model experiments, Transcriptomine contains a small but growing number of small-scale, case-control studies documenting transcriptomic signatures of pathological conditions. To enable the discovery of relationships between NR signaling pathways and disease states, we mapped these data sets to the same parental terms used for biosample (table S2) and model experiment mappings (table S3). In addition, to allow for the identification of cross-talk between NRs and other signaling pathways, a dedicated section of the Regulation Report displays data points from experiments involving small-molecule or genetic manipulation of receptors, enzymes, and transcription factors not related to NR signaling pathways (fig. S3, Other Pathways). Use Case 3 (see below) illustrates how these data points can be used to model potential cross-talk between different signaling pathways.

Drilling down places data points in context and enables their citation

The query form and Regulation Reports are designed to give users the biological “big picture.” To drill down on specific data points, users can first audit results by clicking on a data point in the Regulation Report to display an FCI window (fig. S3) containing the gene symbol, fold change and *P* value, experiment name, biosample, and species. Having identified a data point of interest, the user clicks on the More Information link in the FCI window to display the FCD window (Fig. 1E and fig. S5). The

FCD window contains an Experiment Information section, which details gene manipulations, or small-molecule dose and concentration or duration of treatment. To place the experiment in the context of the larger data set, additional information is provided in the Data Set Information section, which includes the data set name and description fields, which define any gene or small-molecule symbols present in the experiment name and description. To encourage users to provide attribution to the original data set creators, the FCD window provides for one-click download of the data set citation into a user’s reference manager (5). Finally, to close the “logical loop” between the data mining and data set interfaces (Fig. 1F), the FCD window contains a link back to the data set page from which the data point originated.

Use Cases demonstrate linking signaling pathways, gene, and tissues using Transcriptomine

We next describe a series of Use Cases demonstrating the ability of Transcriptomine Regulation Reports to initiate or substantiate research hypotheses that address underappreciated relationships between NR signaling pathways, genes, and tissues. Three of these (Use Cases 1, 2, and 4) involve NR signaling pathways, whereas a fourth (Use Case 3) involves a non-NR pathway to demonstrate the expanding utility of our resource to investigators in other signaling disciplines. Links to the relevant Transcriptomine gene Regulation Reports are embedded in the text, and Use Cases are visually summarized (Fig. 3). The Supplementary Materials contain summaries of all Use Case query parameters and the corresponding gene list downloads (table S3). Transcriptomine data points supporting each Use Case are cited using paired references in which the Transcriptomine data set is immediately followed by its associated research article.

Use Case 1 shows how biology described in an article is indicated by data points in data sets that predated the article but existed in isolation from each other and were therefore difficult to integrate into a coherent biological narrative. The unfolded protein response (UPR) is a highly conserved program of gene expression regulated by multiple signaling pathways that alleviate endoplasmic reticulum stress associated with the accumulation of unfolded proteins (30). A 2015 article in *Oncogene* described how 17 β E2 induced key components of the UPR, including *HSPA5*, *SERP1*, and *XBPI*, in breast cancer cells in an ER α /ESR1-dependent manner (31). The authors postulated that induction by 17 β E2 of this pathway might constitute an adaptive stress response that promoted resistance to tamoxifen therapy, and described it as a novel 17 β E2-ER α /ESR1-regulated pathway. Transcriptomine contained numerous data points from multiple independent, pre-2015 data sets pointing to transcriptional regulation by the ERs and estrogens pathway of components of the UPR pathway, including *HSPA5* (<https://goo.gl/5n3hW6>) [mouse uterus (32–37) and rat vagina (38, 39); data file S1, Query 2], *XBPI* (<https://goo.gl/MijdTty>) [MCF-7 cells (21, 22, 40–45) and mouse testis (46, 47); data file S1, Query 3], and *SERP1* (<https://goo.gl/XwFbPo>) [MCF-7 cells (21, 22) and mouse uterus (32–35); data file S1, Query 4] (Fig. 3, Use Case 1). Collectively, these data points form a reasonable rationale for designing experiments to determine the mechanistic basis for regulation of the UPR by the ERs and estrogens signaling pathway. These queries also indicate regulation of UPR genes by the AR and androgens pathway [*HSPA5* in PC3 (48, 49) and LNCaP (50, 51) cells and mouse liver (52, 53), *XBPI* in mouse testis (46, 47), and *SERP1* in LNCaP cells (54–57) and mouse testis (46, 47)], by the GR and glucocorticoids pathway [*HSPA5* in mouse liver (58, 59) and rat kidney (60, 61) and *SERP1* in mouse bone marrow–derived macrophages (62, 63) and human HLE B-3 cells (64, 65)], and by the PPARs and fatty acids pathway [*XBPI* during 3T3-L1 adipogenesis (66, 67)].

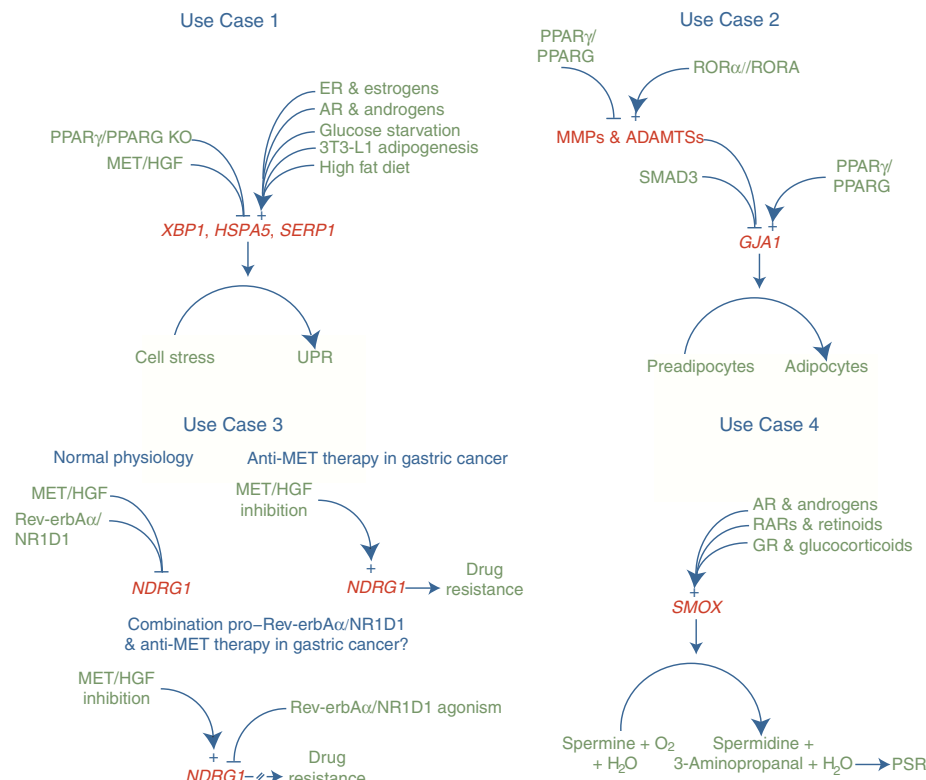


Fig. 3. Use Cases illustrating development of research hypotheses in Transcriptomine. Use Case 1: ER and estrogen signaling pathway regulates the UPR. KO, knockout. Use Case 2: Regulation of 3T3-L1 adipogenesis by the PPAR γ /PPARG and ROR α /RORA signaling pathways involves antagonistic regulation of gap junction formation. Use Case 3: Combination Rev-erbA α /NR1D1 agonism and MET/HGF antagonism in chemoresistant gastric cancer. Use Case 4: The spermine oxidase (SMOX) gene is regulated by multiple NR signaling pathways in different physiological contexts. All mechanistic relationships are inferred from data points in Transcriptomine data sets. Use Case parameters and corresponding search results are contained in data file S1. Direct links to relevant gene Regulation Reports are embedded in the respective sections in the main text.

Use Case 2 focuses on the utility of Transcriptomine Animal and Cell Model experiments in building confidence in a hypothesis previously uncharacterized in the research literature, for the role of two distinct NR signaling pathways in the regulation of 3T3-L1 adipogenesis. Gap junction channels are cellular structures that play important roles in cell-cell communication by enabling the propagation of chemical and electrical signals. Although the literature on the relationship between gap junctions and adipocyte differentiation is conflicting, a consensus appears to be that the presence of gap junctions early during adipogenesis is required for synchronous initiation of the differentiation program (68–70). Although published evidence links NR signaling pathways to regulation of fat cell differentiation (71), there have been no published studies linking such pathways to regulation of gap junction formation during this process. Consistent with the increased expression in white adipose tissue (WAT) of the *GJA1* gene (72), the Transcriptomine Regulation Report for *GJA1* in adipose tissue (<https://goo.gl/HXBLA7>) (data file S1, Query 5) indicated that it is dynamically regulated during 3T3-L1 adipogenesis and that this regulation is mediated in part by PPAR γ /PPARG, a master regulator of adipogenesis (66, 67, 73, 74). Transcriptomine data points reflecting the induction of *Gja1* in WAT of mice lacking *Smad3* (75, 76), a mediator of transforming growth factor- β inhibition of adipocyte differentiation (77), provide further evidence for a putative role for *Gja1* in adipogenesis.

To find further evidence that the PPARs and fatty acids pathway regulates adipogenesis in part by regulating gap junction formation, we next asked Transcriptomine whether other genes involved in gap junction formation were regulated in adipose tissue. Versican (encoded by the *Vcan* gene) modulates gap junction communication in 3T3-L1 cells (78). The Transcriptomine Regulation Report for *VCAN* in adipose tissue (<https://goo.gl/OBCqy9>) (data file S1, Query 6) suggested its regulation early in 3T3-L1 adipogenesis (79, 80) and PPAR γ /PPARG-dependent repression in mature 3T3-L1 adipocytes (66, 67, 73, 74). Of possible translational relevance, data points from clinical data sets highlighted differences in *VCAN* expression in subcutaneous and omental fat from obese Pima Indian subjects relative to lean individuals (81, 82).

An antagonistic relationship exists between gap junctions and members of the extracellular matrix metalloproteinase (MMP) family (83–85). If the emerging hypothesis that *GJA1* is induced during adipogenesis in a PPARs and fatty acids pathway-dependent manner is well founded, we anticipated that genes in the MMP family would be repressed during adipogenesis. Transcriptomine queries showed that five members of the MMP family [*Mmp2* (<https://goo.gl/VF7rn4>), *Mmp3* (<https://goo.gl/Y0tzvn>), *Mmp9* (<https://goo.gl/KmD3p6>), *Mmp14* (<https://goo.gl/my63v1>), and *Mmp23b* (<https://goo.gl/3AF9dx>)] were subject to PPAR γ /Pparg-dependent repression during adipogenesis (Fig. 3, Use Case 2, and data file S1, Queries 7 to 11) (66, 67). Transcriptomine queries also indicated PPAR γ /Pparg-dependent repression during 3T3-L1 adipogenesis of five members of the ADAMTS (a disintegrin and metalloprotease with thrombospondin motif) family of zinc proteases [*Adamts1* (<https://goo.gl/zvLyjm>), *Adamts2* (<https://goo.gl/8fvA8r>), *Adamts4* (<https://goo.gl/4I3DII>), *Adamts5* (<https://goo.gl/Dmz9Se>), and *Adamts10* (<https://goo.gl/rIQJMS>)] (data file S1, Queries 12 to 16) (66, 67, 79, 80). Although gap junction proteins are not known substrates for ADAMTS family members, the overlapping sensitivity of ADAMTS and MMP family members to certain inhibitors (86) suggests that this may be the case. In support of our Transcriptomine-generated hypothesis, two studies unknown to us that were published while our manuscript was in revision (87, 88) reported an antagonistic relationship between *ADAMTS1* and adipogenesis.

A final source of evidence for our hypothesis related to signaling by ROR α /Rora, which inhibits adipocyte differentiation (89, 90). Adipogenesis is also inhibited in mice overexpressing hepatic cholesterol sulfotransferase (91), which catalyzes the formation of cholesterol sulfate, a ROR α /Rora agonist (92). Given these observations, we anticipated that ROR α /Rora signaling would antagonize the expression of genes under the control of the PPARs and fatty acids pathway in adipose tissue. Consistent with this, a total of 10 genes encoding MMP (data file S1, Queries 7 to 11) and ADAMTS (data file S1, Queries 12 to 16) metalloproteinase

family members, as well as *Vcan* (data file S1, Query 6) were shown by Transcriptomine to be repressed in ROR α /RORA-depleted mouse WAT (93, 94).

Use Case 3 illustrates the utility of the direct link to a gene Regulation Report from a gene list on a data set page (Fig. 1B) to extend the scope of a data set beyond its associated research article. It also emphasizes the value of Transcriptomine to investigators of signaling pathways other than those involving NRs. Aberrant activation of the pathway mediated by receptor tyrosine kinase MET and its physiological ligand, hepatocyte growth factor (HGF), occurs in gastric cancer (95–97). Although MET pathway inhibitors have been evaluated as therapeutic options in gastric cancer (95, 98), such strategies are frequently associated with the acquisition of resistance. A 2009 *Science Signaling* article probing the mechanistic basis of this resistance (99) included an expression profiling data set evaluating the transcriptomic response of a panel of human gastric cancer cell lines to the MET inhibitor PHA-665752 (PHA665) (100). A gene that is not discussed in the article but is consistently induced by PHA665 across all experiments in the associated data set is n-Myc downstream-regulated gene 1 (*NDRG1*), which has been implicated in tumorigenesis and metastasis in various cancers (101). The potential role of *NDRG1* in resistance of gastric cancer cells to MET pathway inhibition implicated by these data points is suggested by the role of *NDRG1* in multidrug resistance in neuroblastoma cells (102). Moreover, its stabilization of the epithelial state in nasopharyngeal cancer cells (103) is consistent with the association between resistance to the MET inhibitor KRC-108 and transition to an epithelial phenotype (104). The Transcriptomine *NDRG1* Regulation Report (<https://goo.gl/F76hnk>) (accessed by clicking on the Transcriptomine query link in the FCI window of one of the *NDRG1* data points) highlighted repression of *Ndr1* by the NR Rev-erbA α /Nr1d1 (data file S1, Query 17) (105, 106). These data points are consistent with the transcriptional repression of *NDRG1* by iron (107) and the identity of heme as a Rev-erbA α /NR1D1 agonist (108, 109). Moreover, heme oxygenase, which depletes cellular heme abundance, has been associated in gastric cancer cells with resistance to apoptosis (110), itself a mechanism of cancer drug resistance. Epidemiological evidence indicates an inverse relationship between the incidence of gastric cancer and body iron stores (111). Collectively, the Transcriptomine *NDRG1* data points and the subsequent focused literature mining suggest the hypothesis that repression by Rev-erbA α /NR1D1 signaling of *NDRG1* is a possible basis for combination therapy to circumvent acquisition of resistance to MET inhibitors in gastric cancer (Fig. 3, Use Case 3).

Use Case 4 focuses on spermine oxidase, encoded by the *SMOX* gene, which catalyzes the oxidation of spermine to spermidine, and is an important component of the polyamine stress response (PSR) (Fig. 3, Use Case 4) (112). The Transcriptomine *SMOX* Regulation Report (<https://goo.gl/tQzDxN>) contained data points suggesting its regulation by multiple NR signaling pathways in various biological contexts (Fig. 3, Use Case 4, and data file S1, Query 18). First, the RAR and retinoids signaling section of the *SMOX* Regulation Report contained nearly 30 data points, indicating its induction during a time course study of all-*trans* retinoic acid (ATRA) in HL-60 leukocytes (113, 114). This is in contrast to a conventional literature search, which failed to locate any published studies connecting *SMOX* with the RARs and retinoids signaling pathway. The RARs and retinoids pathway induces terminal differentiation of HL-60 cells (115–117), and spermidine is required for this differentiation (118, 119). The Transcriptomine data points therefore constitute a basis for the hypothesis that ATRA drives terminal differentiation in HL-60 cells in part through induction of *SMOX* to maintain

cellular spermidine concentrations. Second, data points in the *SMOX* Regulation Report reflected its induction by the GR and glucocorticoids signaling pathway and suggested a mechanistic basis for the relationship between glucocorticoids and the PSR (120). Finally, data points in the AR and androgens pathway section of the *SMOX* Regulation Report indicated repression of *SMOX* after disruption of AR signaling in skeletal muscle (121, 122). Given the contribution of androgen signaling to skeletal muscle development (123) and the status of *SMOX* as a marker of myogenic differentiation (124), Transcriptomine helped develop the hypothesis that the promotion of skeletal muscle differentiation by the AR and androgens pathway is mediated in part by induction of *SMOX*.

ChIP-Seq evidence from the Gene Transcription Regulatory Database supports Transcriptomine Use Cases

Finally, we wished to supplement the transcriptomic evidence used to model the signaling pathway–gene relationships described above using cistromic evidence from the Gene Transcriptional Regulation Database (GTRD) (125), a compendium of uniformly processed chromatin immunoprecipitation sequencing (ChIP-Seq) data sets from the National Center for Biotechnology Information (NCBI) Short Read Archive. Of the 15 receptor-gene or gene family relationships modeled in the Use Cases, and for which comparable data sets were available in GTRD (regulation of *GREB1* by PPAR γ /PPARG; regulation of genes encoding UPR pathway members by ER α /ESR1, AR, GR/NR3C1, and PPAR γ /PPARG; regulation of *GJA1* and *VCAN* by PPAR γ /PPARG; regulation of genes encoding MMP and ADAMTS family members by PPAR γ /PPARG and ROR α /RORA; regulation of *NDRG1* by Rev-erbA α /NR1D1; and regulation of *SMOX* by GR/NR3C1, RAR α /RARA, and AR), the GTRD provided evidence of receptor promoter binding sites in 13. Specifically, GTRD queries identified binding sites for PPAR γ /PPARG in the *GREB1* gene (fig. S6A); for ER α /ESR1 in the human *HSPA5* (fig. S6B), *XBPI* (fig. S6C), and *SERP1* (fig. S6D) genes (Fig. 3, Use Case 1); for AR in the *HSPA5* (fig. S6E), *XBPI* (fig. S6F), and *SERP1* (fig. S6G) genes; for GR/NR3C1 in the *HSPA5* (fig. S6H) and *SERP1* (fig. S6I) genes; for PPAR γ /PPARG in the *XBPI* gene (fig. S6J); for PPAR γ /PPARG in the mouse *Gja1* (fig. S6K) and human *VCAN* (fig. S6L), *MMP3* (fig. S6M), *MMP14* (fig. S6N), *ADAMTS1* (fig. S6O), *ADAMTS2* (fig. S6P), *ADAMTS4* (fig. S6Q), and *ADAMTS5* (fig. S6R) genes (Fig. 3, Use Case 2); for Rev-erbA α /Nr1d1 in the mouse *Ndr1* gene (Fig. 3, Use Case 3, and fig. S6S); and for GR/NR3C1 (fig. S6T), RAR α /RARA (fig. S6U), and AR (fig. S6V) in the human *SMOX* gene (Fig. 3, Use Case 4). Although no evidence was found in the GTRD for ROR α /RORA binding sites in genes in the MMP or ADAMTS families, previous reports of DNA binding-independent regulation of gene expression by ROR α /RORA (126) indicate that direct transrepression of PPAR γ /PPARG or another pathway by ROR α /RORA cannot be excluded as a mechanism in these contexts.

DISCUSSION

Although other transcriptomic databases exist (127, 128), Transcriptomine differs from these by organizing the data points in the original data sets into biologically meaningful pathway and biosample classes. Our Use Cases demonstrate the routine but detailed insights into mechanistically underdeveloped aspects of NR signaling biology that can be achieved through reuse of appropriately biocurated transcriptomic data sets. That transcriptomic data sets have biological value beyond the specific contexts in which they were originally generated is not entirely surprising and is not the primary focus of this article. Rather, it is the

ease and intuition with which their underlying data points are made available to the community for reuse, sharing, and citation that represents the true value of Transcriptomine. It should be stressed that Transcriptomine data points do not per se meet the rigorous standards for unambiguously establishing mechanism and require detailed validation in subsequent bench experiments. That said, the reduction in time and effort realized through Transcriptomine in locating, connecting, and interpreting these data points greatly enhances their accessibility and usability by the research community.

Aside from their enhancement of Transcriptomine usability, our biocuration efforts have additional value in shedding light on the relative volumes of data points representing different pathways and physiological systems. Certain experimental paradigms, most notably the study of the ERs and estrogens pathway in MCF-7 cells, are heavily overrepresented among expression array data sets. Such redundancy is not necessarily undesirable because the statistical power of numerous independent data sets helps suppress the technical and biological noise inherent in such experiments. That said, our high-level overview highlights large discrepancies in coverage of particular signaling pathways and physiological systems by 'omics data sets. Given the major public health impact in Western nations of diseases of the renal system, for example, the underrepresentation of archived data sets involving manipulations of the mineralocorticoid receptor (MR) and mineralocorticoid pathway (Fig. 1A), or performed in kidney model systems (classified in Fig. 1B in the Other category), is surprising. Similarly, publically archived 'omics data sets relevant to VDR signaling in bone are in remarkably short supply, whereas others targeting receptors such as TR2 and TR4, for example, are virtually nonexistent. We suggest that the disparity in coverage might benefit from the redistribution of funding to generate and archive reference discovery-scale data sets for signaling pathways that, although currently less well characterized, have the potential to provide important insights into the regulation of physiology by the NR superfamily.

Although recognized as a reliable surrogate for cellular protein abundance (129, 130), global measures of transcript relative abundance constitute only one of the 'omics modalities currently available for interrogating the biology of NR signaling pathways. Data sets documenting the impact of NR pathway manipulations on transcription factor DNA binding events (ChIP-Seq), protein-protein interactions (interactomics), posttranslational modification (PTMomics), and cellular metabolite levels (metabolomics) are being generated in increasing quantities. Equally, NR signaling pathways constitute only a subset of the signal transduction pathways that affect mammalian cells, and investigators studying other pathways have been equally prodigious in their generation of discovery-scale data sets. Adding to the growing volume of 'omics-scale data sets, the clinical research community is generating and making publically available discovery-scale profiles of molecular events accompanying pathogenesis in cancer and other disease states (131). The biocuration and web design principles outlined in this paper are readily abstractable to integration of these diverse classes of 'omics data sets and, pending future funding, will support expansion of our resource into a pan-omics discovery tool for cellular signaling pathways and human disease.

MATERIALS AND METHODS

Designation of a data set

Designation of a group of experiments as a data set occurs on a case-by-case basis after closely studying the associated publication. In general, a

data set is defined as a group of experiments associated with a single PubMed Identifier (PMID) in a single unique biosample. The term biosample was adopted to align our terminology with those currently in use by NCBI (132) and European Bioinformatics Institute (133) and refers here to the tissue or cell type from which the assay starting material (in the case of transcriptomic data sets, mRNA) was derived. The exceptions to this rule are where the data set is specifically designed to compare the transcriptomes of two or more distinct biosamples. On occasion, however, GEO data set depositions contain groups of experiments from two different biosamples that bear no relation to each other in the context of the associated article. In these instances, experiments from distinct biosamples are treated as distinct data sets, with the same PMID.

Data set naming convention

Data set names contain four key elements: (i) the 'omics category of the data set: currently predominantly transcriptomic but positioned for expansion into other 'omics data set categories; (ii) key regulatory molecules (full names spelled out along with familiar and approved symbols, where these exist, and which are used in the data set description and in its constituent experiment names and descriptions); (iii) a brief reference to the organ or organ system—providing either the cell line name or the organ or tissue name for animal studies; and (iv) where appropriate, a reference to specific data set designs (for example, time course or dose dependence).

Data set description convention

Data set descriptions provide the orientation and exposition required for a user to place the experimental results in a clear biological context: This is particularly important in the case of animal and cell models with which the user may not be familiar. In addition to enhancing their discoverability through text-based search engines, the adoption of these semantic standards for data set names and descriptions is designed to ensure a consistent, predictable experience for users of Transcriptomine as they browse from one data set to the next. This is in contrast to the names of the primary archived data sets, which are often simply the name of the associated article, and often give only an impression of the overall experimental design.

Experiment naming convention

Disambiguation of experiment names in the Nuclear Receptor Signaling Atlas (NURSA) transcriptomic data sets is an essential part of the user experience, for example, when experiment names are viewed side by side in the gene list drop-down on individual data set pages or when fold changes are compared in a Transcriptomine scatterplot. Equally, removing redundancy in experiment names makes for a cleaner, less visually cluttered UI. To accommodate the high degree of complexity in experimental design in a consistent human- and machine-readable manner, we developed an experiment naming convention in which up to four defined elements are combined to convey the essential elements of an experimental contrast while making the experiment name unique within a specific data set (table S4). The first element is the regulatory molecule core contrast. Regulatory molecules are small molecules or genes subjected to one or more manipulations, identified using codes (table S5). Genes in the NR signaling field are referred to using familiar symbols and by, with increasing frequency, their approved gene symbols. Accordingly, genes are identified using the name most commonly in use in the field and the species-appropriate approved gene symbol. The second element is the small-molecule treatment concentration/dose and duration. Small molecule-regulated transcriptomes are often compared

across multiple concentrations and doses and durations of treatment, each represented by an individual experiment in a larger data set. In these experiment names, the regulatory molecule is separated by a pipe (|) from the concentration/dose and duration information. The third element relates to molecules on both sides of a contrast. Experimental contrasts are often set up with molecules on both sides of the contrast to provide for comparison of a given differential expression data point between two or more biological contexts. Molecules that appear on both sides of the contrast are indicated in parentheses after the core contrast (and the Concentration/Dose and Duration, if present), along with a manipulation code as required. Multiple molecules are separated by a “+” sign. In the case of data sets comparing gene manipulations with the wild-type state, control experiments are designated (WT) to indicate an animal wild-type for the gene of interest. The fourth element relates to Additional Disambiguators, which appear after a dash (-) at the end of the Experiment Name to make an Experiment Name unique within a Data set. Examples of additional disambiguators include an abbreviated biosample, duration, gender, strain, or some other relevant annotations.

Receptor-ligand biocuration

Where available, we adopted mappings curated by the International Union of Pharmacology Guide to Pharmacology (GtoPdb) resource (134) or DrugBank (135). Where mappings were not available from either of these resources, we created these de novo based on biocuration of the literature on a specific ligand or regulatory small molecule (136).

Biosample (tissue and cell line) biocuration

Biosample biocuration adopts a practical, functional approach that is predicated less upon the anatomical origin of a particular organ or tissue and more on its systemic physiological function of its constituent tissues and cell types. Leukocytes, for example, are located primarily in the cardiovascular compartment, but are almost invariably studied in the context of the immune system, and are classified as such. Biosamples occasionally undergo modifications designed to recapitulate a specific functional or physiological context—for example, treatment of cells with adipogenic (66) or inflammatory (137) chemical cocktails. To suppress complexity and clutter in the query and visualization UIs, such modifications are not encoded in the biosample vocabulary but rather are described in detail in the Experiment-level metadata section of Materials and Methods.

Regulatory molecule unique identifiers

Experiments are mapped to regulatory molecules using unique identifiers for small-molecule (PubChem) or gene (Entrez Gene ID) manipulations. Similarly, all experiments are mapped to unique identifiers for major anatomical and cell ontologies, including UBERON (138), CLO (139), and BRENDA (140).

Experiment descriptions

Experiment descriptions are assigned using a systematic approach similar to that adopted for the description of their parent data set, which provides for a consistent and predictable user experience when browsing between different experiments (fig. S1B).

Experiment Numbers

The assignment of Experiment Numbers to Experiments is an important step in the annotation and curation process because it determines how experiments will be grouped relative to each other in the UI, for example, in the pull-down menus that will be used to browse the indi-

vidual gene lists (fig. S1B). The order of display of experiments is particularly important in data sets documenting the dose/concentration or time dependence of a given small molecule–regulated transcriptome.

NURSA web application

The NURSA hub is a gene-centric Java Enterprise Edition 6, web-based application around which other gene, RNA, protein, drug/ligand, clinical trial, disease, and other data from dozens of external databases are collected. This resource was redesigned to improve the UI and experience, particularly through implementation of responsive design frameworks. Transcriptomine is one component of this resource and links with NURSA Molecule pages, on which gene, RNA, protein, disease, and other data are summarized, and Data Set pages. All software is freely available at www.github.com/BCM-DLDC/nursa.

After undergoing semiautomated processed and biocuration as described above, the data and annotations are stored in NURSA's Oracle 11g database. RESTful web services expose Transcriptomine data, which are served to responsively designed views in the UI, were created using a Flat UI Toolkit with a combination of JavaScript, D3.js, AJAX, HTML5, and CSS3. JavaServer Faces and PrimeFaces are the primary technologies behind the UI. Transcriptomine has been optimized for Firefox 24+, Chrome 30+, Safari 5.1.9+, and Internet Explorer 9+, with validations performed in BrowserStack and load testing in LoadUIWeb. XML describing each data set and experiment is generated and submitted to CrossRef to mint DOIs.

Programmatic access through API

Application programming interface (API) documentation is available on the NURSA website at www.nursa.org/nursa/rs/index.jsf.

Literature searches

All literature searches carried out in developing the Use Cases involved reasonable effort on the part of the curator and employed resources typically used by bench scientists, such as PubMed and Google.

SUPPLEMENTARY MATERIALS

www.sciencesignaling.org/cgi/content/full/10/476/eaah6275/DC1

Fig. S1. The NURSA data set page.

Fig. S2. Transcriptomine query form.

Fig. S3. Transcriptomine *GREB1* Regulation Report (Pathway view).

Fig. S4. Transcriptomine *CA12* Regulation Report (Biosample view).

Fig. S5. FCD window.

Fig. S6. Corroborating receptor promoter binding evidence from the GTRD.

Table S1. Receptor and small-molecule signaling pathway mappings.

Table S2. Cell line and tissue biosample mappings to physiological systems and organs.

Table S3. Examples of animal and cell models and clinical data sets.

Table S4. Experiment naming convention.

Table S5. Nonstandard abbreviations in NURSA experiment names.

Data file S1. Use Case Query parameters and search results.

REFERENCES AND NOTES

1. N. J. McKenna, R. B. Lanz, B. W. O'Malley, Nuclear receptor coregulators: Cellular and molecular biology. *Endocr. Rev.* **20**, 321–344 (1999).
2. D. J. Mangelsdorf, C. Thummel, M. Beato, P. Herrlich, G. Schütz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon, R. M. Evans, The nuclear receptor superfamily: The second decade. *Cell* **83**, 835–839 (1995).
3. N. J. McKenna, Discovery-driven research and bioinformatics in nuclear receptor and coregulator signaling. *Biochim. Biophys. Acta* **1812**, 808–817 (2011).
4. S. A. Ochsner, D. L. Steffen, C. J. Stoeckert Jr., N. J. McKenna, Much room for improvement in deposition rates of expression microarray datasets. *Nat. Methods* **5**, 991 (2008).

5. Y. F. Darlington, A. Naumov, A. McOwiti, W. H. Kankanamge, L. B. Becnel, N. J. McKenna, Improving the discoverability, accessibility and citability of omics datasets: A case report. *J. Am. Med. Inform. Assoc.* **23**, 388–393 (2017).
6. S. A. Ochsner, C. M. Watkins, A. McOwiti, X. Xu, Y. F. Darlington, M. D. Dehart, A. J. Cooney, D. L. Steffen, L. B. Becnel, N. J. McKenna, Transcriptome, a web resource for nuclear receptor signaling transcriptomes. *Physiol. Genomics* **44**, 853–863 (2012).
7. Y. Shang, M. Brown, Molecular determinants for the tissue specificity of SERMs. *Science* **295**, 2465–2468 (2002).
8. M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
9. J. S. Amberger, C. A. Bocchini, F. Schiettecatte, A. F. Scott, A. Hamosh, OMIM.org: Online Mendelian Inheritance in Man (OMIM), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **43**, D789–D798 (2015).
10. J. M. Rae, M. D. Johnson, J. O. Scheys, K. E. Cordero, J. M. Larios, M. E. Lippman, GREB1 is a critical regulator of hormone dependent breast cancer growth. *Breast Cancer Res. Treat.* **92**, 141–149 (2005).
11. J. M. Rae, M. D. Johnson, K. E. Cordero, J. O. Scheys, J. M. Larios, M. M. Gottardis, K. J. Pienta, M. E. Lippman, GREB1 is a novel androgen-regulated gene required for prostate cancer growth. *Prostate* **66**, 886–894 (2006).
12. M. J. Bolt, F. Stossi, J. Y. Newberg, A. Orjalo, H. E. Johansson, M. A. Mancini, Coactivators enable glucocorticoid receptor recruitment to fine-tune estrogen receptor transcriptional responses. *Nucleic Acids Res.* **41**, 4036–4048 (2013).
13. J. S. Carroll, C. A. Meyer, J. Song, W. Li, T. R. Geistlinger, J. Eeckhoutte, A. S. Brodsky, E. K. Keeton, K. C. Fertuck, G. F. Hall, Q. Wang, S. Bekiranov, V. Sementchenko, E. A. Fox, P. A. Silver, T. R. Gingeras, X. S. Liu, M. Brown, Time course analysis of the 17 β -estradiol (17 β E2) dependent transcriptome in MCF-7 breast cancer cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/endTZMGB6Z (2006).
14. J. S. Carroll, C. A. Meyer, J. Song, W. Li, T. R. Geistlinger, J. Eeckhoutte, A. S. Brodsky, E. K. Keeton, K. C. Fertuck, G. F. Hall, Q. Wang, S. Bekiranov, V. Sementchenko, E. A. Fox, P. A. Silver, T. R. Gingeras, X. S. Liu, M. Brown, Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.* **38**, 1289–1297 (2006).
15. E. C. Chang, B. S. Katzenellenbogen, T. H. Charn, J. A. Katzenellenbogen, S. Park, W. G. Helferich, B. Komm, Time course analysis of the 17 β -estradiol (17 β E2)-, genistein (Gen)-, and fulvestrant (Fulv)-regulated, estrogen receptor- β (ER β /ESR2)-dependent transcriptomes in MCF-7 breast cancer cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/JTNKFMkDSu (2008).
16. E. C. Chang, T. H. Cham, S.-H. Park, W. G. Helferich, B. Komm, J. A. Katzenellenbogen, B. S. Katzenellenbogen, Estrogen receptors α and β as determinants of gene expression: Influence of ligand, dose, and chromatin binding. *Mol. Endocrinol.* **22**, 1032–1043 (2008).
17. S. E. Wardell, D. Kazmin, D. P. McDonnell, Comparison of selective estrogen receptor modulator-regulated transcriptomes in MCF-7 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/AxDt4iAOyZ (2012).
18. S. E. Wardell, D. Kazmin, D. P. McDonnell, Research resource: Transcriptional profiling in a cellular model of breast cancer reveals functional and mechanistic differences between clinically relevant SERM and between SERM/estrogen complexes. *Mol. Endocrinol.* **26**, 1235–1248 (2012).
19. T. Miller, C. Arteaga, Analysis of the fulvestrant (Fulv)-dependent transcriptomes in two long-term estrogen-deprived breast cancer cell lines, MCF-7 and HCC1428. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/Hxz5Xa79WN (2011).
20. T. W. Miller, J. M. Balko, E. M. Fox, Z. Ghazoui, A. Dunbier, H. Anderson, M. Dowsett, A. Jiang, R. A. Smith, S.-M. Maira, H. C. Manning, A. M. González-Angulo, G. B. Mills, C. Higham, S. Chanthaphaychith, M. G. Kuba, W. R. Miller, Y. Shyr, C. L. Arteaga, ER α -dependent E2F transcription can mediate resistance to estrogen deprivation in human breast cancer. *Cancer Discov.* **1**, 338–351 (2011).
21. S. Al Saleh, F. Al Mulla, Y. A. Luqmani, Analysis of the estrogen receptor- α (ER α /ESR1)-dependent transcriptome in MCF-7 breast cancer cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/5F49JYoGXU (2011).
22. S. Al Saleh, F. Al Mulla, Y. A. Luqmani, Estrogen receptor silencing induces epithelial to mesenchymal transition in human breast cancer cells. *PLOS ONE* **6**, e20610 (2011).
23. S. Welle, M. Muyan, S. Nott, Y. Huang, Analysis of the estrogen receptor- α (ER α /ESR1)-dependent transcriptome in 17 β -estradiol (17 β E2)-treated MDA-MB-231 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/52SBAzR3nb (2009).
24. S. L. Nott, Y. Huang, X. Li, B. R. Fluharty, X. Qiu, W. V. Welshons, S. Yeh, M. Muyan, Genomic responses from the estrogen-responsive element-dependent signaling pathway mediated by estrogen receptor α are required to elicit cellular alterations. *J. Biol. Chem.* **284**, 15277–15288 (2009).
25. B. W. O'Malley, N. J. McKenna, Coactivators and corepressors: What's in a name? *Mol. Endocrinol.* **22**, 2213–2214 (2008).
26. N. J. McKenna, B. W. O'Malley, Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465–474 (2002).
27. D. Kazmin, S. Gaillard, L. L. Grasfeder, C. Y. Chang, D. P. McDonnell, Transcriptomic analysis of (1) wild type peroxisome proliferator activated receptor- γ coactivator 1 (PGC1 α /PPARGC1); (2) estrogen receptor related receptor- α (ERR α /ESRR α)-disrupted PGC1 α ; and (3) ERR α -specific PGC1 α /PPARGC1 in HepG2 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/datasets.02004 (2006).
28. S. Gaillard, L. L. Grasfeder, C. L. Haeffele, E. K. Lobenhofer, T.-M. Chu, R. Wolfinger, D. Kazmin, T. R. Koves, D. M. Muoio, C.-y. Chang, D. P. McDonnell, Receptor-selective coactivators as tools to define the biology of specific receptor-coactivator pairs. *Mol. Cell* **24**, 797–803 (2006).
29. N. Oksala, M. Levula, M. Peltö-Huikko, L. Kytömäki, J. T. Soini, J. Salenius, M. Kähönen, P. J. Karhunen, R. Laaksonen, S. Parkkila, T. Lehtimäki, Carbonic anhydrases II and XII are up-regulated in osteoclast-like cells in advanced human atherosclerotic plaques—Tampere Vascular Study. *Ann. Med.* **42**, 360–370 (2010).
30. D. Ron, P. Walter, Signal integration in the endoplasmic reticulum unfolded protein response. *Nat. Rev. Mol. Cell Biol.* **8**, 519–529 (2007).
31. N. Andruska, X. Zheng, X. Yang, W. G. Helferich, D. J. Shapiro, Anticipatory estrogen activation of the unfolded protein response is linked to cell proliferation and poor survival in estrogen receptor α -positive breast cancer. *Oncogene* **34**, 3760–3769 (2015).
32. K. S. Korach, S. C. Hewitt, B. J. Deroo, S. F. Grissom, J. B. Collins, Time course transcriptomic analysis of the 17 β -estradiol (17 β E2)-regulated transcriptome in mouse uterus. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/Nk1NOZFLJF (2003).
33. S. C. Hewitt, B. J. Deroo, K. Hansen, J. Collins, S. Grissom, C. A. Afshari, K. S. Korach, Estrogen receptor-dependent genomic responses in the uterus mirror the biphasic physiological response to estrogen. *Mol. Endocrinol.* **17**, 2070–2083 (2003).
34. S. C. Hewitt, K. E. Gerrish, K. S. Korach, Comparative analysis of the 17 β -estradiol (17 β E2)-biphenol A (BPA)- and hydroxychlor (HPTE)-regulated transcriptomes in mouse uterus. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/8hn2WrxOsu (2011).
35. S. C. Hewitt, K. S. Korach, Estrogenic activity of bisphenol A and 2,2-bis(p-hydroxyphenyl)-1,1-trichloroethane (HPTE) demonstrated in mouse uterine gene profiles. *Environ. Health Perspect.* **119**, 63–70 (2011).
36. J. C. Kwekel, L. D. Burgoon, J. W. Burt, J. R. Harkema, T. R. Zacharewski, Time course comparative analysis of the 17 α -ethinylestradiol (17 α EE2)-regulated transcriptome in mouse uterus. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/XwobySFqSi (2005).
37. J. C. Kwekel, L. D. Burgoon, J. W. Burt, J. R. Harkema, T. R. Zacharewski, A cross-species analysis of the rodent uterotrophic program: Elucidation of conserved responses and targets of estrogen signaling. *Physiol. Genomics* **23**, 327–342 (2005).
38. S. A. Jelinsky, Time course analysis of the 17 β -estradiol (17 β E2)-dependent transcriptome in rat vagina. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/Fh5YpJbKia (2008).
39. S. A. Jelinsky, S. E. Choe, J. S. Crabtree, M. M. Cotreau, E. Wilson, K. Saraf, A. J. Dornier, E. L. Brown, B. J. Peano, X. Zhang, R. C. Winneker, H. A. Harris, Molecular analysis of the vaginal response to estrogens in the ovariectomized rat and postmenopausal woman. *BMC Med. Genomics* **1**, 27 (2008).
40. V. Bourdeau, J. Deschênes, D. Laperrière, J. H. White, S. Mader, Analysis of the 17 β -estradiol (17 β E2)-regulated, cycloheximide (CHX)-sensitive transcriptome in MCF-7 breast cancer cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/2bBDIKDoRJ (2007).
41. V. Bourdeau, J. Deschênes, D. Laperrière, M. Aid, J. H. White, S. Mader, Mechanisms of primary and secondary estrogen target gene regulation in breast cancer cells. *Nucleic Acids Res.* **36**, 76–93 (2008).
42. P. Bhat-Nakshatri, G. Wang, H. Appaiah, N. Luktuke, J. S. Carroll, T. R. Geistlinger, M. Brown, S. Badve, Y. Liu, H. Nakshatri, Analysis of the AKT-regulated transcriptome in 17 β -estradiol (17 β E2)-treated MCF-7 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/Gjlg8XmOk (2008).
43. P. Bhat-Nakshatri, G. Wang, H. Appaiah, N. Luktuke, J. S. Carroll, T. R. Geistlinger, M. Brown, S. Badve, Y. Liu, H. Nakshatri, AKT alters genome-wide estrogen receptor α binding and impacts estrogen signaling in breast cancer. *Mol. Cell. Biol.* **28**, 7487–7503 (2008).
44. J. Frasar, Transcriptomic analysis of TNF α - and 17 β -estradiol (17 β E2)-treated MCF-7 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/dErlbHBCBE (2008).
45. J. Frasar, A. Weaver, M. Pradhan, Y. Dai, L. D. Miller, C.-Y. Lin, A. Stanculescu, Positive cross-talk between estrogen receptor and NF- κ B in breast cancer. *Cancer Res.* **69**, 8918–8925 (2009).
46. E. M. Snyder, C. L. Small, Y. Li, M. D. Griswold, Analysis of the testosterone (Test) and 17 β -estradiol (17 β E2)-regulated transcriptomes in mouse efferent ductules and caput epididymis. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/4v3EzdhMx (2009).
47. E. M. Snyder, C. L. Small, Y. Li, M. D. Griswold, Regulation of gene expression by estrogen and testosterone in the proximal mouse reproductive tract. *Biol. Reprod.* **81**, 707–716 (2009).
48. B. Lin, Analysis of the androgen receptor (AR)-dependent transcriptome in PC3 prostatic carcinoma cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/wGxm43bqz (2009).
49. B. Lin, J. Wang, X. Hong, X. Yan, D. Hwang, J.-H. Cho, D. Yi, A. G. Utleg, X. Fang, D. E. Schones, K. Zhao, G. S. Omenn, L. Hood, Integrated expression profiling and

- ChIP-seq analyses of the growth inhibition response program of the androgen receptor. *PLoS ONE* **4**, e6589 (2009).
50. C. E. Massie, A. Lynch, Time course analysis of the R1881-regulated transcriptome in the human LNCaP prostate cancer cell line. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/43PvAd2d8U (2010).
 51. C. E. Massie, A. Lynch, A. Ramos-Montoya, J. Boren, R. Stark, L. Fazli, A. Warren, H. Scott, B. Madhu, N. Sharma, H. Bon, V. Zecchini, D.-M. Smith, G. M. DeNicola, N. Mathews, M. Osborne, J. Hadfield, S. Macarthur, B. Adryan, S. K. Lyons, K. M. Brindle, J. Griffiths, M. E. Gleave, P. S. Rennie, D. E. Neal, I. G. Mills, The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. *EMBO J.* **30**, 2719–2733 (2011).
 52. J. Krücken, D. Delić, C. Weiland, H. Vohr, F. Wunderlich, Analysis of the testosterone (Test)-regulated transcriptome in female mouse liver. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/R3Fq2573TV (2009).
 53. D. Delć, C. Grosser, M. Dkhil, S. Al-Quraishy, F. Wunderlich, Testosterone-induced upregulation of miRNAs in the female mouse liver. *Steroids* **75**, 998–1004 (2010).
 54. J. C. Nwachukwu, P. Mita, R. Ruoff, S. Ha, Q. Wang, S. J. Huang, S. S. Taneja, M. Brown, W. L. Gerald, M. J. Garabedian, S. K. Logan, Analysis of the UXT and R1881 dependent transcriptomes in LNCaP prostate cancer cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/gZmlbnYZ2 (2009).
 55. J. C. Nwachukwu, P. Mita, R. Ruoff, S. Ha, Q. Wang, S. J. Huang, S. S. Taneja, M. Brown, W. L. Gerald, M. J. Garabedian, S. K. Logan, Genome-wide impact of androgen receptor trapped clone-27 loss on androgen-regulated transcription in prostate cancer cells. *Cancer Res.* **69**, 3140–3147 (2009).
 56. H. Hieronymus, T. R. Golub, Analysis of the R1881-regulated transcriptome in gedunin-treated LNCaP cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/vWp6CUBniU (2006).
 57. H. Hieronymus, J. Lamb, K. N. Ross, X. P. Peng, C. Clement, A. Rodina, M. Nieto, J. Du, K. Stegmaier, S. M. Raj, K. N. Maloney, J. Clardy, W. C. Hahn, G. Chiosis, T. R. Golub, Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. *Cancer Cell* **10**, 321–330 (2006).
 58. J. R. Revollo, J. A. Cidlowski, Analysis of the dexamethasone (Dex)-regulated, hairy and enhancer of split-1 (Hes1)-dependent murine hepatic transcriptome. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/mbbl1zihup (2014).
 59. J. R. Revollo, R. H. Oakley, N. Z. Lu, M. Kadmiel, M. Gandhavadi, J. A. Cidlowski, HE51 is a master regulator of glucocorticoid receptor-dependent gene expression. *Sci. Signal.* **6**, ra103 (2013).
 60. R. Almon, Time course analysis of the methylprednisolone (MePred)-regulated transcriptome in rat kidney. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/YaEgBOfWjd (2004).
 61. R. R. Almon, W. Lai, D. C. DuBois, W. J. Jusko, Corticosteroid-regulated genes in rat kidney: Mining time series array data. *Am. J. Physiol. Endocrinol. Metab.* **289**, E870–E882 (2005).
 62. R. T. Yu, N. H. Uhlenhaut, R. M. Evans, Analysis of the dexamethasone (Dex)-regulated transcriptome in lipopolysaccharide (LPS)-activated inflammatory mouse macrophages. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/NjyEDq96T7 (2013).
 63. N. H. Uhlenhaut, G. D. Barish, R. T. Yu, M. Downes, M. Karunasingh, C. Liddle, P. Schwalie, N. Hübner, R. M. Evans, Insights into negative regulation by the glucocorticoid receptor from genome-wide profiling of inflammatory cistromes. *Mol. Cell* **49**, 158–171 (2013).
 64. V. K. Gupta, A. Galante, P. Soteropoulos, B. J. Wagner, Analysis of the dexamethasone (Dex) dependent transcriptome in immortalized human lens epithelial cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/Oulg4ukCds (2005).
 65. V. Gupta, A. Galante, P. Soteropoulos, S. Guo, B. J. Wagner, Global gene profiling reveals novel glucocorticoid induced changes in gene expression of human lens epithelial cells. *Mol. Vis.* **11**, 1018–1040 (2005).
 66. M. Schupp, A. G. Cristancho, M. I. Lefterova, E. A. Hanniman, E. R. Briggs, D. J. Steger, M. Qatanani, J. C. Curtin, J. Schug, S. A. Ochsner, N. J. McKenna, M. A. Lazar, Analysis of the peroxisome proliferator-activated receptor- γ (PPAR γ /Pparg)-regulated transcriptome during adipogenesis in mouse 3T3-L1 adipocytes. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/nWjv1UhrEE (2009).
 67. M. Schupp, A. G. Cristancho, M. I. Lefterova, E. A. Hanniman, E. R. Briggs, D. J. Steger, M. Qatanani, J. C. Curtin, J. Schug, S. A. Ochsner, N. J. McKenna, M. A. Lazar, Re-expression of GATA2 cooperates with peroxisome proliferator-activated receptor- γ depletion to revert the adipocyte phenotype. *J. Biol. Chem.* **284**, 9458–9464 (2009).
 68. T. Yanagiya, A. Tanabe, K. Hotta, Gap-junctional communication is required for mitotic clonal expansion during adipogenesis. *Obesity* **15**, 572–582 (2007).
 69. A. Yeganeh, G. L. Stelmack, R. R. Fandrich, A. J. Halayko, E. Kardami, P. Zahradka, Connexin 43 phosphorylation and degradation are required for adipogenesis. *Biochim. Biophys. Acta* **1823**, 1731–1744 (2012).
 70. B. D. Pope, C. R. Warren, K. K. Parker, C. A. Cowan, Microenvironmental control of adipocyte fate and function. *Trends Cell Biol.* **26**, 745–755 (2016).
 71. E. D. Rosen, C. J. Walkey, P. Puigserver, B. M. Spiegelman, Transcriptional regulation of adipogenesis. *Genes Dev.* **14**, 1293–1307 (2000).
 72. S. Burke, F. Nagajyothi, M. M. Thi, M. Hanani, P. E. Scherer, H. B. Tanowitz, D. C. Spray, Adipocytes in both brown and white adipose tissue of adult mice are functionally connected via gap junctions: Implications for Chagas disease. *Microbes Infect.* **16**, 893–901 (2014).
 73. J. Rong, V. Zhang, Time course transcriptomic analysis of rosiglitazone (Rosi)-treated 3T3-L1 adipocytes. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/rP497GvEr2 (2009).
 74. J. X. Rong, J. L. Klein, Y. Qiu, M. Xie, J. H. Johnson, K. M. Waters, V. Zhang, J. A. Kashatus, K. S. Remlinger, N. Bing, R. M. Crosby, T. K. Jackson, S. M. Witherspoon, J. T. Moore, T. E. Ryan, S. D. Neill, J. C. Strum, Rosiglitazone induces mitochondrial biogenesis in differentiated murine 3T3-L1 and C3H/10T1/2 adipocytes. *PPAR Res.* **2011**, 179454 (2011).
 75. H. Yadav, C. Quijano, A. K. Kamaraju, O. Gavrilova, R. Malek, W. Chen, P. Zervas, D. Zhigang, E. C. Wright, C. Steulten, P. Sun, S. Lonning, M. Skarulis, A. E. Sumner, T. Finke, S. G. Rane, Analysis of the Tgfb1/Smad3-regulated white adipose tissue (WAT) transcriptome in a mouse model of diet induced obesity. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/yG7RrbFzAc (2015).
 76. H. Yadav, C. Quijano, A. K. Kamaraju, O. Gavrilova, R. Malek, W. Chen, P. Zervas, D. Zhigang, E. C. Wright, C. Steulten, P. Sun, S. Lonning, M. Skarulis, A. E. Sumner, T. Finkel, S. G. Rane, Protection from obesity and diabetes by blockade of TGF- β /Smad3 signaling. *Cell Metab.* **14**, 67–79 (2011).
 77. L. Choy, R. Derynck, Transforming growth factor- β inhibits adipocyte differentiation by Smad3 interacting with CCAAT/enhancer-binding protein (C/EBP) and repressing C/EBP transactivation function. *J. Biol. Chem.* **278**, 9609–9619 (2003).
 78. W. Sheng, H. Dong, D. Y. Lee, W.-y. Lu, B. B. Yang, Versican modulates gap junction intercellular communication. *J. Cell. Physiol.* **211**, 213–219 (2007).
 79. M. S. Hamza, S. Pott, V. B. Vega, J. S. Thomsen, G. S. Kandhadayar, K. P. Chiu, C. L. Wei, Y. Ruan, E. T. Liu, Analysis of the peroxisome proliferator-activated receptor- γ (PPAR γ /Pparg)-regulated adipogenic transcriptome in 3T3-L1 adipocytes. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/4XYkPBMymD (2009).
 80. M. S. Hamza, S. Pott, V. B. Vega, J. S. Thomsen, G. S. Kandhadayar, P. W. Ng, K. P. Chiu, S. Pettersson, C. L. Wei, Y. Ruan, E. T. Liu, De-novo identification of PPAR γ /RXR binding sites and direct targets during adipogenesis. *PLoS ONE* **4**, e4907 (2009).
 81. Y. Lee, S. Nair, E. Rousseau, D. B. Allison, G. P. Page, P. Tataranni, C. Bogardus, P. A. Permana, Analysis of the obese transcriptomic signature in subcutaneous white adipose tissue (WAT) in Pima Indians. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/8Uyh5lBaft (2005).
 82. Y. H. Lee, S. Nair, E. Rousseau, D. B. Allison, G. P. Page, P. A. Tataranni, C. Bogardus, P. A. Permana, Microarray profiling of isolated abdominal subcutaneous adipocytes from obese vs non-obese Pima Indians: Increased expression of inflammation-related genes. *Diabetologia* **48**, 1776–1783 (2005).
 83. M. De Bock, N. Wang, E. Decrock, G. Bultynck, L. Leybaert, Intracellular cleavage of the Cx43 C-terminal domain by matrix-metalloproteinases: A novel contributor to inflammation? *Mediators Inflamm.* **2015**, 257471 (2015).
 84. X. Wu, W. Huang, G. Luo, L. A. Alain, Hypoxia induces connexin 43 dysregulation by modulating matrix metalloproteinases via MAPK signaling. *Mol. Cell. Biochem.* **384**, 155–162 (2013).
 85. M. L. Lindsey, G. P. Escobar, R. Mukherjee, D. K. Goshorn, N. J. Sheats, J. A. Bruce, I. M. Mains, J. K. Hendrick, K. W. Hewett, R. G. Gourdie, L. M. Matrisian, F. G. Spinale, Matrix metalloproteinase-7 affects connexin-43 levels, electrical conduction, and survival after myocardial infarction. *Circulation* **113**, 2919–2928 (2006).
 86. M. L. Moss, J. M. White, M. H. Lambert, R. C. Andrews, TACE and other ADAM proteases as targets for drug discovery. *Drug Discov. Today* **6**, 417–426 (2001).
 87. J. C. Wong, K. C. Krueger, M. J. Costa, A. Aggarwal, H. Du, T. L. McLaughlin, B. J. Feldman, A glucocorticoid- and diet-responsive pathway toggles adipocyte precursor cell activity in vivo. *Sci. Signal.* **9**, ra103 (2016).
 88. S.-Z. Chen, L.-F. Ning, X. Xu, W.-Y. Jiang, C. Xing, W.-P. Jia, X.-L. Chen, Q.-Q. Tang, H.-Y. Huang, The miR-181d-regulated metalloproteinase Adamts1 enzymatically impairs adipogenesis via ECM remodeling. *Cell Death Differ.* **23**, 1778–1791 (2016).
 89. H. Duez, C. Duhem, S. Laitinen, P. S. Patole, M. Abdelkarim, B. Bois-Joyeux, J.-L. Danan, B. Staels, Inhibition of adipocyte differentiation by ROR α . *FEBS Lett.* **583**, 2031–2036 (2009).
 90. N. Ohoka, S. Kato, Y. Takahashi, H. Hayashi, R. Sato, The orphan nuclear receptor ROR α restrains adipocyte differentiation through a reduction of C/EBP β activity and perilipin gene expression. *Mol. Endocrinol.* **23**, 759–771 (2009).
 91. X. Shi, Q. Cheng, L. Xu, J. Yan, M. Jiang, J. He, M. Xu, M. Stefanovic-Racic, I. Sipula, R. M. O'Doherty, S. Ren, W. Xie, Cholesterol sulfate and cholesterol sulfotransferase inhibit gluconeogenesis by targeting hepatocyte nuclear factor 4 α . *Mol. Cell. Biol.* **34**, 485–497 (2014).
 92. J. Kallen, J.-M. Schlaepfli, F. Bitsch, I. Delhon, B. Fournier, Crystal structure of the human ROR α Ligand binding domain in complex with cholesterol sulfate at 2.2 Å. *J. Biol. Chem.* **279**, 14033–14038 (2004).

93. H. S. Kang, K. Okamoto, A. M. Jetten, Analysis of the retinoid orphan receptor- α (ROR α /Rora)-regulated transcriptome in mouse liver and white adipose tissue. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/b2n1yyAn8P (2011).
94. H. S. Kang, K. Okamoto, Y. Takeda, J. Y. Beak, K. Gerrish, C. D. Bortner, L. M. DeGraff, T. Wada, W. Xie, A. M. Jetten, Transcriptional profiling reveals a role for ROR α in regulating gene expression in obesity-associated inflammation and hepatic steatosis. *Physiol. Genomics* **43**, 818–828 (2011).
95. D. L. Jardim, D. de Melo Gagliato, G. S. Falchook, F. Janku, R. Zinner, J. J. Wheler, V. Subbiah, S. A. Piha-Paul, S. Fu, M. B. Murphy, J. Ajani, C. Tang, K. Hess, S. R. Hamilton, S. Roy-Chowdhuri, R. Kurzrock, F. Meric-Bernstam, D. S. Hong, MET aberrations and c-MET inhibitors in patients with gastric and esophageal cancers in a phase I unit. *Oncotarget* **5**, 1837–1845 (2014).
96. G. Minuti, L. Landi, MET deregulation in breast cancer. *Ann. Transl. Med.* **3**, 181 (2015).
97. T. Hara, A. Ooi, M. Kobayashi, M. Mai, K. Yanagihara, I. Nakanishi, Amplification of c-myc, K-sam, and c-met in gastric cancers: Detection by fluorescence in situ hybridization. *Lab. Invest.* **78**, 1143–1153 (1998).
98. E. Gherardi, W. Birchmeier, C. Birchmeier, G. Vande Woude, Targeting MET in cancer: Rationale and progress. *Nat. Rev. Cancer* **12**, 89–103 (2012).
99. A. Z. Lai, S. Cory, H. Zhao, M. Gigoux, A. Monast, M.-C. Guiot, S. Huang, A. Tofigh, C. Thompson, M. Naujokas, V. A. Marcus, N. Bertos, B. Sehat, R. M. Perera, E. S. Bell, B. D. G. Page, P. T. Gunning, L. E. Ferri, M. Hallett, M. Park, Dynamic reprogramming of signaling upon met inhibition reveals a mechanism of drug resistance in gastric cancer. *Sci. Signal.* **7**, ra38 (2014).
100. A. Z. Lai, M. Park, Time-course transcriptomic analysis of a panel of MET receptor tyrosine kinase-inhibited gastric cancer cell lines. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/f11TWZJRA (2014).
101. T. P. Ellen, Q. Ke, P. Zhang, M. Costa, NDRG1, a growth and cancer related gene: Regulation of gene expression and function in normal and disease states. *Carcinogenesis* **29**, 2–8 (2008).
102. D. Zhang, J. Jia, G. Zhao, M. Yue, H. Yang, J. Wang, NDRG1 promotes the multidrug resistance of neuroblastoma cells with upregulated expression of drug resistant proteins. *Biomed. Pharmacother.* **76**, 46–51 (2015).
103. Z.-Y. Hu, W.-B. Xie, F. Yang, L.-W. Xiao, X.-Y. Wang, S.-Y. Chen, Z.-G. Li, NDRG1 attenuates epithelial–mesenchymal transition of nasopharyngeal cancer cells via blocking Smad2 signaling. *Biochim. Biophys. Acta* **1852**, 1876–1886 (2015).
104. D. C. Kim, K. R. Park, Y. J. Jeong, H. Yoon, M. J. Ahn, G. J. Rho, J. Lee, Y. D. Gong, S. Y. Han, Resistance to the c-Met inhibitor KRC-108 induces the epithelial transition of gastric cancer cells. *Oncol. Lett.* **11**, 991–997 (2016).
105. G. Le Martelot, T. Claudel, D. Gatfield, O. Schaad, B. Kornmann, G. Lo Sasso, A. Moschetta, U. Schibler, Analysis of the Rev-Erba/Nr1d1 dependent transcriptome in mouse liver. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/jNzqntK6Gj (2009).
106. G. Le Martelot, T. Claudel, D. Gatfield, O. Schaad, B. Kornmann, G. Lo Sasso, A. Moschetta, U. Schibler, REV-ERB α participates in circadian SREBP signaling and bile acid homeostasis. *PLoS Biol.* **7**, e1000181 (2009).
107. X. Xu, R. Sutak, D. R. Richardson, Iron chelation by clinically relevant anthracyclines: Alteration in expression of iron-regulated genes and atypical changes in intracellular iron distribution and trafficking. *Mol. Pharmacol.* **73**, 833–844 (2008).
108. S. Raghuram, K. R. Stayrook, P. Huang, P. M. Rogers, A. K. Nosie, D. B. McClure, L. L. Burris, S. Khorasanizadeh, T. P. Burris, F. Rastinejad, Identification of heme as the ligand for the orphan nuclear receptors REV-ERB α and REV-ERB β . *Nat. Struct. Mol. Biol.* **14**, 1207–1213 (2007).
109. L. Yin, N. Wu, J. C. Curtin, M. Qatanani, N. R. Szewergold, R. A. Reid, G. M. Waitt, D. J. Parks, K. H. Pearce, G. B. Wisely, M. A. Lazar, Rev-erba, a heme sensor that coordinates metabolic and circadian pathways. *Science* **318**, 1786–1789 (2007).
110. Z.-M. Liu, G. G. Chen, E. K. Ng, W.-K. Leung, J. J. Sung, S. C. Chung, Upregulation of heme oxygenase-1 and p21 confers resistance to apoptosis in human gastric cancer cells. *Oncogene* **23**, 503–513 (2004).
111. A. Fonseca-Nunes, A. Agudo, N. Aranda, V. Arija, A. J. Cross, E. Molina, M. J. Sanchez, H. B. Bueno-de-Mesquita, P. Siersema, E. Weiderpass, V. Krogh, A. Mattiello, R. Tumino, C. Saieva, A. Naccarati, B. Ohlsson, K. Sjöberg, M.-C. Boutron-Ruault, C. Cadeau, G. Fagherazzi, H. Boeing, A. Steffen, T. Kühn, V. Katske, A. Tjønneland, A. Olsen, K.-T. Khaw, N. Wareham, T. Key, Y. Lu, E. Riboli, P. H. Peeters, D. Gavrilu, M. Dorronsoro, J. R. Quirós, A. Barricarte, M. Jenab, R. Zamora-Ros, H. Freisling, A. Trichopoulos, P. Lagiou, C. Bamia, P. Jakszyn, Body iron status and gastric cancer risk in the EURGAST study. *Int. J. Cancer* **137**, 2904–2914 (2015).
112. G. M. Gilad, V. H. Gilad, Overview of the brain polyamine-stress-response: Regulation, development, and modulation by lithium and role in cell survival. *Cell. Mol. Neurobiol.* **23**, 637–649 (2003).
113. S. Huang, G. Eichler, Y. Bar-Yam, D. E. Ingber, Time course transcriptomic analysis of all-trans retinoic acid (ATRA)-treated HL-60 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/fDZIQ8DLV5 (2005).
114. S. Huang, G. Eichler, Y. Bar-Yam, D. E. Ingber, Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys. Rev. Lett.* **94**, 128701 (2005).
115. T. R. Breitman, S. E. Selonick, S. J. Collins, Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2936–2940 (1980).
116. P. A. Meyer, C. Kleinschmitz, Retinoic acid induced differentiation and commitment in HL-60 cells. *Environ. Health Perspect.* **88**, 179–182 (1990).
117. S. J. Collins, K. A. Robertson, L. Mueller, Retinoic acid-induced granulocytic differentiation of HL-60 myeloid leukemia cells is mediated directly through the retinoic acid receptor (RAR- α). *Mol. Cell. Biol.* **10**, 2154–2163 (1990).
118. M. Sugiura, T. Shafman, T. Mitchell, J. Griffin, D. Kufe, Involvement of spermidine in proliferation and differentiation of human promyelocytic leukemia cells. *Blood* **63**, 1153–1158 (1984).
119. D. W. Kufe, J. Griffin, T. Mitchell, T. Shafman, Polyamine requirements for induction of HL-60 promyelocyte differentiation by leukocyte-conditioned medium and phorbol ester. *Cancer Res.* **44**, 4281–4284 (1984).
120. V. H. Gilad, J. M. Rabey, Y. Kimiagar, G. M. Gilad, The polyamine stress response: Tissue-, endocrine-, and developmental-dependent regulation. *Biochem. Pharmacol.* **61**, 207–213 (2001).
121. K. Mo, J. T. Westwood, D. A. Monks, Mutational analysis of the androgen receptor (AR/Ar)-regulated transcriptome in mouse skeletal muscle models of Kennedy disease/spinal bulbar muscular atrophy. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/BuE2FppUS9 (2010).
122. K. Mo, Z. Razak, P. Rao, Z. Yu, H. Adachi, M. Katsuno, G. Sobue, A. P. Lieberman, J. T. Westwood, D. A. Monks, Microarray analysis of gene expression by skeletal muscle of three mouse models of Kennedy disease/spinal bulbar muscular atrophy. *PLoS ONE* **5**, e12922 (2010).
123. A. D. Mooradian, J. E. Morley, S. G. Korenman, Biological actions of androgens. *Endocr. Rev.* **8**, 1–28 (1987).
124. M. Cervelli, E. Fratini, R. Amendola, M. Bianchi, E. Signori, E. Ferraro, A. Lisi, R. Federico, L. Marocci, P. Mariottini, Increased spermine oxidase (SMO) activity as a novel differentiation marker of myogenic C2C12 cells. *Int. J. Biochem. Cell Biol.* **41**, 934–944 (2009).
125. I. Yevshin, R. Sharipov, T. Valeev, A. Kel, F. Kolpakov, GTRD: A database of transcription factor binding sites identified by ChIP-seq experiments. *Nucleic Acids Res.* **45**, D61–D67 (2017).
126. J. M. Lee, I. S. Kim, H. Kim, J. S. Lee, K. Kim, H. Y. Yim, J. Jeong, J. H. Kim, J.-Y. Kim, H. Lee, S.-B. Seo, H. Kim, M. G. Rosenfeld, K. I. Kim, S. H. Baek, ROR α attenuates Wnt/ β -catenin signaling by PKC α -dependent phosphorylation in colon cancer. *Mol. Cell* **37**, 183–195 (2010).
127. A. P. Davis, T. C. Wieggers, B. L. King, J. Wieggers, C. J. Grondin, D. Sciaky, R. J. Johnson, C. J. Mattingly, Generating gene ontology-disease inferences to explore mechanisms of human disease at the comparative toxicogenomics database. *PLoS ONE* **11**, e0155530 (2016).
128. J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J.-P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, T. R. Golub, The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
129. J. J. Li, M. D. Biggin, Gene expression. Statistics quantitates the central dogma. *Science* **347**, 1066–1067 (2015).
130. J. J. Li, P. J. Bickel, M. D. Biggin, System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ* **2**, e270 (2014).
131. J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).
132. National Center for Biotechnology Information (2016).
133. European Bioinformatics Institute (2016).
134. C. Southan, J. L. Sharman, H. E. Benson, E. Faccenda, A. J. Pawson, S. P. H. Alexander, O. P. Buneman, A. P. Davenport, J. C. McGrath, J. A. Peters, M. Spedding, W. A. Catterall, D. Fabbro, J. A. Davies; NC-IUPHAR, The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: Towards curated quantitative interactions between 1300 protein targets and 6000 ligands. *Nucleic Acids Res.* **44**, D1054–D1068 (2016).
135. V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. Maciejewski, D. Arndt, M. Wilson, V. Neveu, A. Tang, G. Gabriel, C. Ly, S. Adamjee, Z. T. Dame, B. Han, Y. Zhou, D. S. Wishart, DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **42**, D1091–D1097 (2014).
136. L. B. Becnel, Y. F. Darlington, S. A. Ochsner, J. R. Easton-Marks, C. M. Watkins, A. McOwiti, W. H. Kankanamge, M. W. Wise, M. DeHart, R. N. Margolis, N. J. McKenna, Nuclear receptor signaling atlas: Opening access to the biology of nuclear receptor signaling pathways. *PLoS ONE* **10**, e0135615 (2015).
137. P. A. Burke, Z. Wang, E. P. Bishop, Analysis of the hepatocyte nuclear factor 4 receptor α (HNF4 α /HNF4A)-regulated transcriptome in inflammatory human HepG2 cells. *Nucl. Recept. Signal. Atlas Datasets* 10.1621/CWhdLJd86i (2009).

138. M. A. Haendel, J. P. Balhoff, F. B. Bastian, D. C. Blackburn, J. A. Blake, Y. Bradford, A. Comte, W. M. Dahdul, T. A. Dececchi, R. E. Druzinsky, T. F. Hayamizu, N. Ibrahim, S. E. Lewis, P. M. Mabee, A. Niknejad, M. Robinson-Rechavi, P. C. Sereno, C. J. Mungall, Unification of multi-species vertebrate anatomy ontologies for comparative biology in Uberon. *J. Biomed. Semantics* **5**, 21 (2014).
139. S. Sarntivijai, Y. Lin, Z. Xiang, T. F. Meehan, A. D. Diehl, U. D. Vempati, S. C. Schürer, C. Pang, J. Malone, H. Parkinson, Y. Liu, T. Takatsuki, K. Saijo, H. Masuya, Y. Nakamura, M. H. Brush, M. A. Haendel, J. Zheng, C. J. Stoeckert, B. Peters, C. J. Mungall, T. E. Carey, D. J. States, B. D. Athey, Y. He, CLO: The cell line ontology. *J. Biomed. Semantics* **5**, 37 (2014).
140. A. Chang, I. Schomburg, S. Placzek, L. Jeske, M. Ulbrich, M. Xiao, C. W. Sensen, D. Schomburg, BRENDA in 2015: Exciting developments in its 25th year of existence. *Nucleic Acids Res.* **43**, D439–D446 (2015).
141. C. Rochette-Egly, Nuclear receptors: Integration of multiple signalling pathways through phosphorylation. *Cell. Signal.* **15**, 355–366 (2003).
142. Y. Igarashi, N. Nakatsu, T. Yamashita, A. Ono, Y. Ohno, T. Urushidani, H. Yamada, Open TG-GATEs: A large-scale toxicogenomics database. *Nucleic Acids Res.* **43**, D921–D927 (2015).

Acknowledgments: We thank current members of the NURSA Beta Testing Group and appreciate valuable discussions with D. Steffen. We thank the past and present NURSA Program staff, R. Margolis and C. Silva [National Institute of Diabetes and Digestive and Kidney

Diseases (NIDDK)] and K. Yoshinaga [National Institute of Child Health and Human Development (NICHD)]. We also thank A. Means and S. Mullican for critical reading of the manuscript. **Funding:** NURSA is supported by grants from the NIDDK (DK097748), with supplemental funding from the Eunice Kennedy Shriver NICHD, and the NIH Big Data to Knowledge (BD2K) program (DK097748-S1) support from the Biomedical Informatics Group of the Biostatistics and Informatics Shared Resource of the Dan L. Duncan Cancer Center, funded by National Cancer Institute grant P30CA125123. **Author contributions:** L.B.B. led the database and web development team. S.A.O. led the biocuration team. Y.F.D. was the day-to-day project manager. W.H.K. developed the UI. A.M. and A.N. were the software architects. M.D. was the systems administrator. N.J.M. and S.A.O. created the concept of Transcriptome. L.B.B., S.A.O., Y.F.D., and N.J.M. designed the project. L.B.B. and N.J.M. wrote the manuscript. N.J.M. led the project. **Competing interests:** The authors declare that they have no competing interests.

Submitted 29 July 2016

Accepted 4 April 2017

Published 25 April 2017

10.1126/scisignal.aah6275

Citation: L. B. Becnel, S. A. Ochsner, Y. F. Darlington, A. McOwiti, W. H. Kankanamge, M. Dehart, A. Naumov, N. J. McKenna, Discovering relationships between nuclear receptor signaling pathways, genes, and tissues in Transcriptome. *Sci. Signal.* **10**, eaah6275 (2017).

Abstracts

One-sentence summary: An updated data-mining web tool enables researchers to understand how nuclear receptors affect genes or cellular processes of interest.

Editor's Summary:

Data mining to understand nuclear receptor signaling

Transcriptomic data are potentially useful for generating mechanistic hypotheses beyond the original experiment in which they were generated and for independently validating unrelated studies. However, data are often generated and presented in disparate contexts and formats, making it difficult to draw connections between different researchers' findings. Becnel *et al.* provide an updated version of Transcriptomine, a data-mining web tool that focuses on nuclear receptor pathway data sets. This tool has been redesigned to be easily used by bench scientists to access and complement data from the published scientific literature. The resource curates more than 500 data sets to allow users to cross-reference information about how different genetic or pharmacological manipulations affect gene expression in different organs or physiological systems and to visualize pathway-gene-tissue relationships. The approach used by the authors can be expanded to other pathways and types of omics data sets.