MINIREVIEW

Research Resources for Nuclear Receptor Signaling Pathways

Neil J. McKenna

Department of Molecular and Cellular Biology and Nuclear Receptor Signaling Atlas Bioinformatics Resource, Baylor College of Medicine, Houston, Texas

Received February 11, 2016; accepted May 19, 2016

ABSTRACT

Nuclear receptor (NR) signaling pathways impact cellular function in a broad variety of tissues in both normal physiology and disease states. The complex tissue-specific biology of these pathways is an enduring impediment to the development of clinical NR small-molecule modulators that combine therapeutically desirable effects in specific target tissues with suppression of off-target effects in other tissues. Supporting the important primary research in this area is a variety of web-based resources that assist researchers in gaining an appreciation of the molecular determinants of the pharmacology of a NR pathway in a given tissue. In this study, selected representative examples of these tools are reviewed, along with discussions on how current and future generations of tools might optimally adapt to the future of NR signaling research.

Introduction

The Nuclear Receptor Superfamily

The 48 proteins of the nuclear receptor (NR) superfamily function as ligand-dependent transcription factors for a diverse set of fat-soluble hormones, vitamins, and dietary lipids (Mangelsdorf et al., 1995). Included in this family are receptors for endocrine steroids (i.e., corticosteroids, progesterone, androgens, and estrogens), fat-soluble vitamins A and D, thyroid hormone, fatty acids, oxysterols, bile acids, and numerous environmental endocrine-disrupting chemicals and xenobiotics. Additional members of this family are referred to as orphan receptors because their ligands remain uncharacterized. As directly druggable regulators of gene expression, NRs and their transcriptional coregulators (Glass et al., 1997; McKenna et al., 1999) are pharmacologically prominent targets for the development of small-molecule therapeutics in a variety of inflammatory, neoplastic, and metabolic conditions (Glass and Ogawa, 2006; Schulman, 2010; Safe et al., 2014).

Biology of NR Signaling Pathways

Signaling pathways involving NRs, their cognate physiologic ligands, and coregulators coordinate the organ- and tissue-specific expression of genes across diverse physiologic systems. Processes regulated by NR signaling pathways include mammalian embryonic development [retinoic acid receptor and all-trans retinoic acid pathway (Mark et al., 2009)]; reproduction [estrogen, progesterone, and androgen receptor pathways (Carpenter and Korach, 2006; Zhao et al., 2008; Rubel et al., 2010)]; metabolism and the inflammatory response [glucocorticoid receptor and peroxisome proliferator-activated receptor (PPAR) subfamily pathways (Pyper et al., 2010; Giordano Attianese and Desvergne, 2015; Granner et al., 2015; Janani and Ranjitha Kumari, 2015)]; and the immune system and bone homeostasis [vitamin D receptor pathway (Christakos et al., 2016)]. Although a full discussion of the biology of NR coregulators is beyond the scope of this minireview, significant findings in this area are the roles of nuclear corepressors 1 and 2, and Mediator 1, and members of the steroid receptor coactivator family in embryonic development, the cardiovascular system, metabolism, and reproduction (McKenna et al., 1999; Giudici et al., 2016). Although NR signaling pathways are commonly named for their principal receptors, ligands and coregulators are key regulatory nodes, and the mechanism by which each pathway communicates the afferent physiologic signal varies between distinct tissues and cell types.

Clinical Pharmacology of NR Signaling Pathways

The extensive biologic footprint of NR signaling pathways is reflected in the intense interest they command as drug targets in a wide variety of human diseases and disorders. The clinical pharmacological agents that target NRs—popularly known as selective receptor modulators—selectively agonize or
antagonize their cognate receptors in a tissue-, cell type-, and promoter-specific manner [comprehensively reviewed by Burris et al. (2013)]. Selective estrogen receptor modulators have found clinical application in estrogen receptor (ER)-positive [tamoxifen (Burris et al., 2013)] and metastatic [toremifene (Mustonen et al., 2014)] breast cancer, osteoporosis [raloxifene (Gizzo et al., 2013)], and vaginal atrophy [lasofoxifene (Pinkerton and Stanczyk, 2014)]. Given their robust antagonism of these signaling conduits in cells mediating the immune and inflammatory responses—B cells, T cells, and macrophages—a variety of glucocorticoid receptor-specific selective receptor modulators is in active clinical use for inflammatory and allergic conditions of the respiratory system (e.g., asthma, rhinitis) and skin (acne, psoriasis) and autoimmune disorders (rheumatoid arthritis), and to suppress local inflammatory responses in musculoskeletal injuries (Burris et al., 2013). The best-characterized—and most controversial—selective modulators of PPAR-γ are the thiazolidinediones, including rosiglitazone, pioglitazone, and troglitazone, which have been used as insulin-sensitizing hypoglycemic agents in the treatment of type 2 diabetes (Soccio et al., 2014). The undesirable side effects of selective estrogen receptor modulators, such as the increased risk of endometrial cancer associated with Tam use (Burris et al., 2013); incidents of heart failure, bone fracture, weight gain, and liver dysfunction associated with selective peroxisome proliferator receptor modulators (Burris et al., 2013); and the effects of glucocorticoid receptor-specific selective receptor modulators on fluid retention, weight gain, and hypertension (Burris et al., 2013), are a signal reminder of the highly nuanced and contextual nature of NR signaling pathway pharmacology.

Research Resources for Analysis of NR Signaling Pathways

Over the past decade, the field of NR signaling has generated a large volume of global datasets that collectively describe sequences of NR and coregulator genes (genomics); the regulation by NRs and coregulators of gene networks in specific target tissues (transcriptomics); protein-protein interactions and post-translational modifications required for the efficient function of NRs and coregulators (proteomics); specific sites of action of NRs in target gene promoters (cistromics); covalent modification of chromatin (epigenomics); and, more recently, their effects on serum and cellular levels of key metabolites and metabolic intermediates (metabolomics) (Fig. 1). Complementing the efforts of the cell biology community in these areas has been the output of the highly active field of clinical chemistry, which has generated a large number of small molecules to probe the fine details of NR signaling pathway function. A greater appreciation of the tissue-specific pharmacology of NR signaling pathways can be assisted by the availability of web-based tools, free or subscription fee-based, that can be routinely accessed by bench scientists with little or no specialist informatics training. We review below a group of examples of such tools, emphasizing where possible their utility for the pharmacology community. For purposes of comparison, we have defined “signaling pathway” broadly, to encompass the following: metabolism of physiologic and synthetic NR ligands; NR and Cora genes, their expression and their protein products; proteomics, including interactions and post-translational modifications; and selected functional endpoints of NR signaling as described by transcriptomics, genomic DNA-binding analysis, and metabolomics. It is not the intent of this review to critically evaluate each resource or point out shortcomings, but rather to highlight those aspects of each resource that we consider to be most useful to the bench pharmacologist. Table 1 contains URLs and literature references for all of the resources cited in the text. Note that although only one of the resources below—the Nuclear Receptor Signaling Atlas (NURSA)—is a curated NR-centric resource, they all encompass information of relevance to NR pathways.

### Ligands

**General Physicochemical Properties.** A number of excellent general chemical resources document general properties of bioactive small molecules, the most comprehensive being PubChem, ChEBI, and DrugBank. Of particular interest to pharmacology field is the detailed information in these resources on absorption, distribution, and excretion of physiologic NR ligands and the safety and toxicity profiles of synthetic analogs and mimetics. DrugBank contains a particularly comprehensive listing of commercially available forms.
Web-based resources for exploring NR signaling pathways

**Research Resources for NR Signaling Pathways**

**Ligand Biosynthesis, Metabolism, and Pharmacogenomics.** The bioavailability of physiologic and synthetic ER agonists and antagonists is determined in larger part by their cellular and systemic concentrations. Information on the physiologic ligand biosynthetic pathways, encompasses the metabolites, the enzymes, and the genes that encode them, as well as their catabolism, are the manually annotated KEGG, HMDB, and DrugBank resources. A particularly attractive aspect of KEGG is its comprehensive graphical depictions of pathways that highlight the relationships between physiologic NR ligands that go a long way to helping the user understand the key molecular interactions and relationships of these molecules. The ovarian steroidogenesis pathway, for example, which encompasses 17βE2 biosynthesis, is displayed as a visual schematic, with the various biosynthetic intermediates and enzymes represented as nodes that link to contextual information, including information on small-molecule inhibitors of those enzymes. Pharmacogenomic interactions between NR ligands that are approved regulatory drugs and single-nucleotide polymorphisms in human genes encoding their catabolic enzymes are the compass of the Pharmacogenomics KnowledgeBase. Like many of the pre-eminent web-based resources in the field, PharmGKB is based in large part upon manual curation.

**NRs and Ligands**

The impact of a given small-molecule regulator of NR function on any given NR signaling pathway is defined in part by their affinities for a range of potential NR-binding partners in a tissue. Although a substantial body of literature has been devoted to this discipline, few sites exist to distill these numerous studies into a researcher-accessible form. The most comprehensive public resource, and the only one in existence that actively curates NR-ligand mappings on a consistent basis, is the International Union of Pharmacology’s *Guide to Pharmacology*. Ligands and receptor mappings, along with essential kinetic information and literature citations, can be found in records in either category of molecule. Mappings of small-molecule NR agonists and antagonists to their cognate receptors can also be found in KEGG and HMDB. Small-molecule perturbants and probes for specific NR signaling pathways, along with information on the assays in which they were screened, are available through the National Institutes of Health’s Molecular Libraries Program.

**NRs and Coregulators**

**Genes, Transcripts, and Proteins.** Another important component of NR signaling pharmacology pathway is the spatiotemporal availability of cognate receptors for small-molecule perturbants. Numerous broad-based gene and protein-centric resources exist that compile, with varying degrees of comprehensive and annotation quality, information on genes, their expression, and the proteins they encode, including National Center for Biotechnology Information Entrez Gene, Ensembl, UniProt, and GeneCards. Of these, GeneCards is, in the author’s opinion, the most comprehensively curated with respect to the various mechanistic, biologic, and clinical aspects of different genes and proteins.

**Expression.** A number of resources exist that contain well-curated, systematically collected datasets documenting tissue-specific gene expression patterns of NRs and coregulators (CoRs). Useful expression profiling-based tools for broadly

### Table 1

<table>
<thead>
<tr>
<th>Resource</th>
<th>Name</th>
<th>URL</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligands</td>
<td>Physicochemical properties; biosynthesis, transport, and catabolism</td>
<td>pubchem.ncbi.nlm.nih.gov, drugbank.ca, hmdb.ca, genome.jp, gene.ncl.nlm.nih.gov</td>
<td>(Kim et al., 2016; Law et al., 2014; Kaneshia et al., 2014; Wishart et al., 2013)</td>
</tr>
<tr>
<td>Ligand pharmacogenomics</td>
<td>IUPHAR GloP</td>
<td>guidetopharmacology.org</td>
<td>(Whirl-Carrillo et al., 2012)</td>
</tr>
<tr>
<td>NRs and ligands</td>
<td>NIH Molecular Libraries</td>
<td><a href="http://mli.nih.gov/mli">http://mli.nih.gov/mli</a></td>
<td>(Southan et al., 2016)</td>
</tr>
<tr>
<td>NRs and CoRs</td>
<td>NCBI Entrez Gene</td>
<td>gene.ncl.nlm.nih.gov</td>
<td>(Schreiber et al., 2015)</td>
</tr>
<tr>
<td>Expression</td>
<td>Ensembl</td>
<td>ensembl.org</td>
<td>(Hubbard et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>UniProt</td>
<td>uniprot.org</td>
<td>(UniProt Consortium, 2014)</td>
</tr>
<tr>
<td></td>
<td>GeneCards</td>
<td>genecards.org</td>
<td>(Safran et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>BioGPS</td>
<td>biogps.org</td>
<td>(Siragusa et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>Allen Brain Atlas</td>
<td>brain-map.org</td>
<td>(Kuan et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>Edinburgh Mouse Atlas</td>
<td>emouseatlas.org ebi.ac.uk/gxa</td>
<td>(Richardson et al., 2015; Petryszak et al., 2016)</td>
</tr>
<tr>
<td></td>
<td>Expression Atlas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteomics</td>
<td>PDB</td>
<td>rcsb.org</td>
<td>(Rose et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>BioGRID</td>
<td>biogrid.org</td>
<td>(Chatr-Aryamontri et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>Pathway Commons</td>
<td>pathwaycommons.org</td>
<td>(Cerami et al., 2011)</td>
</tr>
<tr>
<td>Transcriptomics, cistromics, and epigenomics</td>
<td>Phosphosite Plus</td>
<td>phosphosite.org</td>
<td>(Hornbeck et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>CTDBase</td>
<td>ctd.org</td>
<td>(Davis et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>DAVID</td>
<td>david.ncifcrf.gov</td>
<td>(Sherman et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>NURSA Transcriptomine</td>
<td>nursa.org/ transcriptomine</td>
<td>(Becnel et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>GSEA/msIGDB</td>
<td>broadinstitute.org/gsea/msigdb</td>
<td>(Liberzon et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>Ingenuity</td>
<td>ingenuity.com</td>
<td>(<a href="http://www.ingenuity.com">www.ingenuity.com</a>)</td>
</tr>
<tr>
<td></td>
<td>NextBio</td>
<td>nextbio.com</td>
<td>(<a href="http://www.nextbio.com">www.nextbio.com</a>)</td>
</tr>
</tbody>
</table>
identifying potentially pharmacologically relevant tissues for a given receptor include BioGPS, which takes a specific human, mouse, or rat gene name and returns its relative expression profile across a variety of major tissue types and organs, as well as National Center for Biotechnology Information's Gene Expression Omnibus and European Bioinformatics Institute's Expression Atlas. More granular, anatomical resources include Allen Brain Atlas, Human Protein Atlas, and the Edinburgh Mouse Atlas. Allen Brain Atlas in particular represents an impressive undertaking in both the breadth and depth of its coverage and curation, and mapping of its content to gene symbols provides for easier linking with external resources.

**Proteomics.** An important aspect of the pharmacology of NR signaling pathways is crosstalk between these pathways and other cell signaling pathways. The molecular events associated with such crosstalk are the purview of proteomics, encompassing protein-protein interactions between NRs, other transcription factors and CoRs, and post-translational modifications of these proteins. Probably the most comprehensive resource in existence for protein interactions is BioGRID, which aggregates information from both high-throughput, discovery-driven datasets as well as low-throughput, hypothesis-driven research articles. A search for estrogen receptor α returned a total of more than 1200 physical interactions extracted from more than 330 publications, categorized according to the original assay method, including coimmunoprecipitation/Western, two-hybrid, and native complex reconstitution. Where available, crystal structures for NR, ligand, and/or CoR interactions are available at the Protein Data Bank, which can be searched either by protein name or by small-molecule perturbant, where one is present in the structure. Protein Data Bank features an attractive user interface and highly detailed curation, and, for many publishers, deposition of crystal coordinates in Protein Data Bank is required as a condition of publication of an article. The pre-eminent resource for post-translational modifications of NRs and CoRs and their coregulators is the manually curated Phosphosite Plus, which documents experimentally demonstrated protein phosphorylation and other post-translational modifications, the conservation of these sites across different species, putative targeting pathways, and links to supporting articles in the literature. Finally, the visually appealing Pathway Commons returns network diagrams indicating relationships between estrogen receptor α and other proteins or genes, curated from the literature.

**Transcriptomics, Cistromics, and Epigenomics**

The tissue-selective pharmacology of NR signaling pathways is perhaps best understood in terms of the disparity in events downstream of ligand-receptor interactions, principally the following: 1) genomic binding sites of ligand-receptor complexes and 2) regulation of mRNA transcript levels. Transcriptomic analysis of NR signaling pathways involves global-scale relative abundance studies of these events in response to a specific perturbation, such as ligand versus Veh, knockdown, or knockout versus Con or over-expression of a given receptor. The NURSA Transcriptomine database aggregates NR transcriptomic datasets from public archives, supplements their annotation, and organizes them into gene regulation reference libraries for each of the major NR signaling pathways. Access to these libraries is either through individual datasets, linked where possible to their associated primary research articles, or through the Transcriptomine search engine, which allows for customized queries encompassing NR signaling pathway and organ or tissue. Individual data points in a given contrast are linked back to the complete gene list on the dataset page, where related datasets can be discovered, and the dataset can be cited. The ligand/NR/coregulator-gene-tissue/cell line relationships contained in Transcriptomine allow for evidence gathering, hypothesis generation, and model testing by the bench biologist from any background, and assumes no prior familiarity with the field on their part. Examples of datasets associated with articles in *Molecular Pharmacology* include a study comparing the transcriptomic effects in mouse liver of synthetic agonists of PPAR-α/PPARA (Tijet et al., 2006a,b) and the endocrine-disrupting chemical 2,3,7,8-tetrachlorodibenzo-p-dioxin (Kane et al., 2009a,b), an arylhydrocarbon receptor agonist implicated in transcriptional activation of the constitutive androstane receptor (Petrick and Klaassen, 2007).

A similar aggregative approach is adopted by the Comparative Toxicogenomics Database and NextBio. Like NURSA Transcriptomine, Comparative Toxicogenomics Database (CTD) and NextBio assume no prior knowledge of a specific pathway on the part of a user, allowing them to retrieve information concerning the role of NR signaling pathways on regulation of specific gene or group of genes of interest or, conversely, the top regulated genes for a given pathway. Another category of transcriptomic databases, exemplified by Ingenuity Pathway Analysis, ENRICHR, DAVID, and GSEA/MsigDB, is aimed at more sophisticated informatics users. These resources rank the functional similarity of a user-supplied gene list against a set of reference gene lists compiled from public transcriptomic archives to give the user a sense of what signaling pathways are impacted in their perturbations. Broadly speaking, these two categories of tools are complementary, the latter allowing users to gain a broader perspective of the number of pathways impacted in their experimental models, and the former enabling them to drill down on specific genes and pathways in specific tissues. The ENCODE project—the former based upon de novo datasets and the latter on public archives—compiles genome-wide DNA binding and histone modification (chromatin immunoprecipitation) datasets into searchable resources, where DNA-binding transcription factors and source material (cell line or tissue) can be searched and compared with histone modification patterns at specific promoters.

**Hypothesis Generation Use Cases**

Many data points in discovery-scale or ‘omics datasets are not described in their associated research articles and, as such, are not optimally exposed to search engines such as Google or PubMed or PubMed Central. Accordingly, research resources aggregate and/or annotate these datasets and make them available for data validation and the generation of mechanistic hypotheses in areas of biology that may be new or unfamiliar to their users. In this work, with specific reference to articles in *Molecular Pharmacology*, we illustrate the use of these resources to create or validate connections between distinct signaling pathways in specific physiologic and pathologic contexts.
Citing decreased expression of the mismatch repair gene \textit{MSH2} in oxidatively stressed renal carcinoma cells, Ponnusamy et al. (2016) posited loss of mismatch repair as a potential mechanism for acquired resistance to doxorubicin-induced cytotoxicity in these cells. A search for \textit{MSH2} in the Transcriptome, CTDBase, and NextBio resources found multiple data points documenting its previously uncharacterized repression in liver cells by the xenobiotic phenobarbital (Lambert et al., 2009a,b) and by the GR agonists dexamethasone (Revollo et al., 2013) and methylprednisolone (https://www.nursa.org/nursa/datasets/dataset.jsf?doi=10.1621/SnRGxa6Pwe). In contrast, literature searches of SCOPUS, PubMed, and Google failed to identify a relationship between \textit{MSH2} and these molecules. These data points support the mechanistic hypothesis that the reported suppression of hepatic apoptosis by phenobarbital (Sanders and Thorgersson, 1999) and glucocorticoids (Bailly-Maitre et al., 2001; Oh et al., 2006; Gruver-Yates and Cidlowski, 2013) may be attributable, at least in part, to downregulation of \textit{MSH2} expression. Furthermore, activation by dexamethasone of PXR (Pascussi et al., 2000), which in turn is a well-characterized target of phenobarbital (Willson and Klierwer, 2002), suggests positive crosstalk between these pathways in the liver, which is again consistent with their discordant patterns of regulation of \textit{MSH2}.

A second use case concerns a \textit{Molecular Pharmacology} article in which Jung et al. postulated that induction of the xenobiotic efflux pump \textit{ABCG2} gene by the c-MET/Pi3K pathway played a role in the development of chemoresistance in ovarian cancer cells (Jung, 2015 #96). A search for Transcriptome echoed these findings, showing that \textit{ABCG2} was downregulated in gastric cancer cells treated with a c-Met/HGF inhibitor, PHA665-772. In addition to these corroborating data points, Transcriptome provided evidence postulating previously uncharacterized relationships between \textit{ABCG2} and NR signaling pathways. Repression of \textit{ABCG2} by the AR/androgen signaling pathway in prostate epithelium LNCaP cells (Hieronymus et al., 2006; Kazmin et al., 2006; Nickols and Dervan, 2007; Nwachukwu et al., 2009; Patki et al., 2013) is consistent with its repression by dihydrotestosterone in breast cancer cells (Chua et al., 2016) and suggests a possible mechanism for the relatively low epithelial expression levels of \textit{ABCG2} compared with other prostate cell types (Pascal et al., 2007). Transcriptome also provided evidence for induction of \textit{ABCG2} by 17\(\beta\)-estradiol in osteoblasts (Monroe et al., 2005; Cvoro et al., 2008; Krum et al., 2008; Ball et al., 2009; Paruthiyil et al., 2009). Induction of \textit{ABCG2} in bone is consistent with recent reports of the support by 17\(\beta\)E2 of the osteogenic lineage in a variety of stem cell populations (Taskiran and Evren, 2011; Irmak et al., 2014; Li et al., 2014). Moreover, \textit{ABCG2} expression was also induced in bone by diarylpropionitrile, which is selective for the ER\(\beta\)/ESR2, the ER subform that predominates in bone (Paruthiyil et al., 2009). The effects of the ER/estrogen signaling pathway on \textit{ABCG2} expression in the bone contrast with its more familiar repression of \textit{ABCG2} in mammary gland experimental model systems (Imai et al., 2005) and suggest the testable hypothesis that induction of \textit{ABCG2} by the ER/estrogen signaling pathway supports the osteogenic lineage. These examples serve to illustrate the confidence that results when numerous independent datasets cross-validate to postulate a specific gene tissue-signaling pathway relationship.

**Concluding Remarks**

A broad variety of databases, knowledge bases, and tools exists to support the efforts of bench researchers in the field of NR signaling and its related disciplines. The best of these combine research-focused user interfaces, robust manual curation, and full attribution and acknowledgment of the original studies and their authors. Reviewing them, however, it is difficult to escape the conclusion that they could offer the user so much more if there were better integration between them, such that scientists in one discipline could be readily exposed to information curated from another. This is certainly a complex task, but, given the increasing investment by funding agencies in the infrastructure to support the management of biomedical data, the opportunity is greater now than it has been before. Meaningful integration will require databases to adopt common standards for the exchange of their data, and funding agencies are best placed to ensure that such standards are adopted. These agencies should also ensure that funds are prioritized to support those tools that are both easily locatable by researchers and useful to them. Modern social media seems well placed to allow researchers to provide real-time information on the user base of the various resources—reviews such as this, for example, would be more informative if their authors had access to objective metrics of the impact and utility of these resources in their respective fields. Moreover, improved linkages between journal articles—still the primary means by which researchers consume scientific information—and tools for analyzing the underlying datasets would go a long way toward raising awareness of the number and diversity of resources available to researchers.

As the question mark in Fig. 1 suggests, a notable deficit, for NR and cytoplasmic signaling pathways alike, is the absence of web tools for analyzing their tissue-specific impact on cellular metabolites in normal physiology and metabolic disease. Despite the sharp rise in recent years in signaling metabolomic studies—the number of such studies has increased by more than 1700% over the past decade, compared with an overall growth of signaling articles in PubMed of 75% over the same period—there is currently no freely available resource in which the regulation of cellular metabolism by signal transduction pathways can be compared and contrasted in a user-friendly fashion. Although standards for metabolomics data do exist, albeit in nascent form, deposition of these datasets is yet to be mandated by funding bodies. A brief survey of the recent literature, the details of which are beyond the scope of this review, determined that the deposition of metabolomics datasets in the NR signaling field in public archives is the exception rather than the rule. Given the widely accepted potential of metabolomics to bridge the gap between cell signaling and translational therapeutics (Sreekumar et al., 2009; Hirschev et al., 2010), this seems a missed opportunity for the research community, and there is a collective responsibility on the part of publishers, data repository sponsors, and investigators to redress this situation.

The feudal nature of scientific research and communication—investigators, publishers, funding organizations, citation managers, databases, knowledge bases—complicates attempts to bring order to the often overwhelming number of distinct sources of biomedical data to which the bench researcher is exposed. This complexity has most likely supported the perception held by some that investment in the
infrastructure to support data reuse has been misspent and might be better allocated to hypothesis-driven research. In the author's own experience, however, the deficits in the utility of many tools are attributable largely to the existence in the public domain of many improperly archived and poorly annotated datasets. To address this long-standing issue, funding agencies should consider supporting resources that provide assistance to investigators in the deposition of their datasets in repositories, so as not to burden their research with the relatively mundane, but important, task of depositing the datasets. Equally, community knowledge bases and data repositories should map their curated content to controlled vocabularies and ontologies to support automated and federated distribution, so that their content is visible and leveragable across diverse research communities. Sharing of data holds great promise, but bench scientists will fully embrace data reuse only when those data are freely and easily accessible, comprehensively and accurately annotated, and intuitively presented and integrated with other similar resources.

The field of NR signaling looks forward to a new generation of biomedical research resources based on genuine and enduring commitments to these principles.

Authorship Contributions

Wrote or contributed to the writing of the manuscript: McKenna.

References


