

Article

A Sub-Pathway Based Method to Identify Candidate Agents for Ankylosing Spondylitis

Kai Chen [†], Yingchuan Zhao [†], Yu Chen, Chuanfeng Wang, Ziqiang Chen, Yushu Bai, Xiaodong Zhu and Ming Li ^{*}

Department of Orthopedics, Changhai Hospital, Shanghai 200433, China

[†] These authors contributed equally to this work.

^{*} Author to whom correspondence should be addressed; E-Mail: minglilimingming@hotmail.com; Tel./Fax: +86-021-3116-6666.

Received: 26 June 2012; in revised form: 25 September 2012 / Accepted: 26 September 2012 /

Published: 22 October 2012

Abstract: The need for new therapeutics for Ankylosing Spondylitis (AS) is highlighted by the general lack of efficacy for most agents currently available for this disease. Many recent studies have detailed molecular pathways in AS, and several molecule-targeting agents are undergoing evaluation. We aimed to explore the mechanism of AS and identify biologically active small molecules capable of targeting the sub-pathways which were dysregulated in the development of AS. By using the GSE25101 microarray data accessible from the Gene Expression Omnibus database, we first identified the differentially expressed genes (DEGs) between AS samples and healthy controls, followed by the sub-pathway enrichment analysis of the DEGs. In addition, we propose the use of an approach based on targeting sub-pathways to identify potential agents for AS. A total of 3,280 genes were identified as being significantly different between patients and controls with p -values < 0.1 . Our study showed that neurotrophic signaling pathway and some immune-associated pathways may be involved in the development of AS. Besides, our bioinformatics analysis revealed a total of 15 small molecules which may play a role in perturbing the development of AS. Our study proposes the use of an approach based on targeting sub-pathways to identify potential agents for AS. Candidate agents identified by our approach may provide the groundwork for a combination therapy approach for AS.

Keywords: Ankylosing Spondylitis; bioinformatics; connectivity map; drug discovery

1. Introduction

Ankylosing Spondylitis (AS) is a chronic inflammatory arthritis involving primarily the sacroiliac joints and the axial skeleton that can lead to bone resorption and bone formation, ultimately resulting in ankylosis [1,2]. AS can cause inflammation or injury to other joints away from the spine, as well as to other organs, such as the eyes, heart, lungs, and kidneys. AS is the prototype of spondyloarthropathies (SpA), a related family of disorders with common clinical features. Some 30 years ago, the strong association between human leukocyte antigen B27 (HLA-B27) and susceptibility to AS was reported, however, HLA-B27 accounts only for about 5% of the genetic risk in AS [3]. Non-HLA-B27 genetic and environmental factors have an important role in its development and progression. Genome-wide association studies have identified several other non-HLA susceptibility genes such as IL23R and ERAP1 in AS [4,5].

The therapeutic management of AS includes regular physical exercise together with the use of NSAIDs. Second-line treatments, such as sulfasalazine, are required in cases of NSAID-refractory AS [6]. Besides, TNF α blockers have been shown to be the most promising treatment, helping many patients receive a significant reduction, though not elimination, of their inflammation and pain. Several cytokine inhibitor agents are currently being evaluated as agents for AS including tocilizumab and rituximab [7,8]. Although there has been interest in the evaluation of several of the currently available targeted therapies for use in AS, the rational use of these agents has been hindered by the paucity of knowledge about the cellular mechanisms involved in response to therapy. Consequently there remains a need for more effective therapies, or for approaches that can improve therapeutic responses to the treatment of AS.

In this present study, we sought to explore the mechanism of AS, and then identify biologically active small molecules capable of targeting the sub-pathways which were dysregulated in the development of AS. Candidate agents identified by our approach may provide the groundwork for a combination therapy approach for AS. However, further evaluations for their potential use in the treatment of AS are still needed.

2. Results

2.1. Differentially Expressed Genes Analysis Between AS and Healthy Controls

To identify the genes specifically differentially expressed between AS patients and controls, we obtained the publically available microarray dataset GSE25101 from the GEO database and carried out the classical t-test corrected for multiple comparisons. A total of 3,280 genes were considered significantly differentially expressed with $p < 0.1$.

2.2. Sub-Pathway Enrichment Analysis

We used Subpathwayminer to annotate these differentially expressed genes to entire metabolic pathways and sub-pathways ($k = 3$). With the strict cutoff of p less than 0.01, our system identified 35 enriched sub-pathways corresponding to 18 entire metabolic pathways (Table 1). In fact, all 18 of the entire pathways can be divided into two groups: Immune-associated or cancer associated. These 35

sub-pathways closely interact with each other in their corresponding complete pathways and were significantly dysregulated in the development of AS.

Table 1. The enriched entire pathways and sub-pathways (k = 3).

Entire pathway ID	Entire pathway name	Sub-pathway ID	p-value
path:00520	Amino sugar and nucleotide sugar metabolism	path:00520_1	0.003405
path:04662	B cell receptor signaling pathway	path:04662_2	0.008457
		path:04110_16	0.009995
		path:04110_21	0.001077
path:04110	Cell cycle	path:04110_26	0.000558
		path:04110_3	0.002999
		path:04110_4	0.000804
path:04062	Chemokine signaling pathway	path:04062_20	0.006018
		path:04664_10	0.006018
path:04664	Fc epsilon RI signaling pathway	path:04664_3	0.006018
		path:04664_5	0.003016
		path:04664_9	0.008457
path:04666	Fc gamma R-mediated phagocytosis	path:04666_3	0.008109
		path:04666_4	0.006903
path:04510	Focal adhesion	path:04510_26	0.002765
path:00052	Galactose metabolism	path:00052_7	0.001238
		path:05160_1	0.000137
path:05160	Hepatitis C	path:05160_4	2.81E-05
		path:05160_5	0.000643
path:04650	Natural killer cell mediated cytotoxicity	path:04650_3	0.004229
		path:04722_18	0.002847
		path:04722_21	0.001121
path:04722	Neurotrophin signaling pathway	path:04722_22	0.000804
		path:04722_23	0.009995
		path:04722_5	0.003948
path:04621	NOD-like receptor signaling pathway	path:04621_1	0.005705
path:05223	Non-small cell lung cancer	path:05223_6	0.008457
path:04330	Notch signaling pathway	path:04330_1	0.00757
		path:05200_10	0.008404
path:05200	Pathways in cancer	path:05200_18	0.008457
		path:05200_51	0.007638
		path:05200_52	0.002695
path:04810	Regulation of actin cytoskeleton	path:04810_31	0.000643
path:05222	Small cell lung cancer	path:05222_1	0.001288
path:04620	Toll-like receptor signaling pathway	path:04620_6	0.004707

2.3. Identification of Candidate Small Molecules

According to the gene-expression profiles data from 6,100 treatment-control pairs (instances) involving 1,309 bioactive small molecules in the Connectivity Map, we performed DEGs analysis for these 1,309 small molecules. A total of 1,221 out of the 1,309 small molecules have DEGs. For the

1,221 small molecules, we then used SubpathwayMiner to annotate these DEGs to entire metabolic pathways and sub-pathways. With the strict cutoff of FDR less than 0.005, our system identified 233 enriched sub-pathways corresponding to 107 small molecules.

2.4. Network Construction between Sub-Pathways and Small Molecules

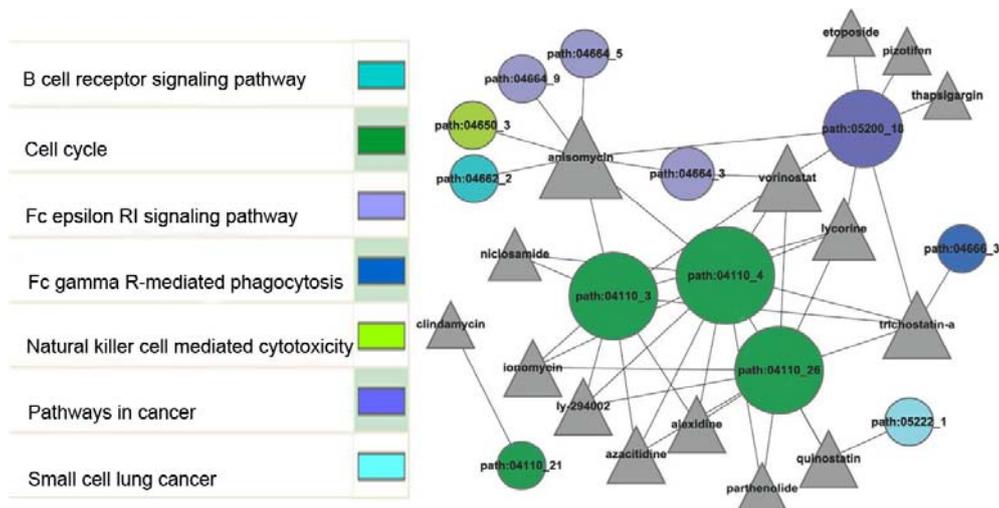
By integrating the previous 35 sub-pathways and these 233 sub-pathways, we could select the significant overlapping sub-pathways, that is, these sub-pathways are related to both AS and small molecules. Thus, we could hypothesize that these small molecules may play a role in perturbing the development of AS. A total of 15 small molecules were identified (Table 2).

Table 2. The significant overlapping small molecules.

Drug bank ID	Small molecule	<i>p</i> -value	Number of overlaps	Type
DB07374	anisomycin	3.34E-11	8	experimental
DB02546	vorinostat	4.98E-08	5	approved
—	quinostatin	6.28E-08	3	
—	lycorine	1.79E-06	4	
—	alexidine	3.47E-06	3	
—	ionomycin	3.47E-06	3	
—	ly-294002	3.47E-06	3	
—	trichostatin A	3.99E-06	5	
—	azacitidine	5.20E-06	3	
DB06803	niclosamide	0.000754	2	approved
DB06803	parthenolide	0.003105	2	approved
DB01190	clindamycin	0.012813	1	approved
—	pizotifen	0.025472	1	
—	thapsigargin	0.029659	1	
DB00773	etoposide	0.046238	1	approved

We then built a network between the small molecules and sub-pathways which were perturbed by the small molecules by integrating the relationships above (Figure 1). In this network, some small molecules can perturb one or several sub-pathways by themselves, while some small molecules perturb sub-pathways by cooperating with other small molecules.

Figure 1. Small molecules perturb the sub-pathways in AS. The triangular nodes stand for small molecules and the circular nodes stand for sub-pathways. The sub-pathways in same color are included in same entire pathways. The description of the corresponding entire pathways is shown on the left part.



3. Discussion

Gene expression profiling in disease reveals the underlying gene activity changes contributing to the disease and enables identification of targets for therapeutic intervention. In this study, we used the gene expression profile downloaded from GEO to explore the mechanism of AS development. Besides, we showed the utility of using bioinformatics analysis for the identification of new therapeutics for AS. A total of 3,280 genes were identified as being differentially expressed between AS patients and healthy controls. Sub-pathway mining result showed 35 sub-pathways corresponding to 18 entire pathways were dysregulated in AS, and a total of 15 small molecules were identified which may play a role in perturbing the development of AS.

Current approaches typically study entire pathways, whether by singular enrichment analysis or by gene set enrichment analysis [9]. However, sub-pathways analysis may be more suitable than entire pathways for identification of drug responses [10]. In this study, we identified 35 sub-pathways from 18 KEGG pathways. A number of these sub-pathways have well-documented inflammatory roles or an action on bone/cartilage metabolism, such as sub-pathways associated with immune and amino sugar and nucleotide sugar metabolism. Interestingly, the sub-pathway enrichment analysis indicated a dysregulated “neurotrophin signaling pathway” in the development of AS. This result agrees with a previous study which suggested the up-regulation of neurotrophins and their receptors in spondyloarthritis [11] and is also consistent with the report that the expression of neurotrophins and their receptors is related to inflammation and may be involved in the local disease processes [12].

Neurotrophins are a family of trophic factors involved in differentiation and survival of neural cells [13]. The neurotrophin family consists of nerve growth factor (NGF), brain derived neurotrophic factor (BDNF), neurotrophin 3 (NT-3), and neurotrophin 4 (NT-4) [14]. Neurotrophins are trophic factors whose actions are crucial for differentiation, survival and the organization of neuronal connections during postnatal development [15,16]. Neurotrophin signaling is regulated by connecting a variety of

intracellular signaling cascades, which include MAPK pathway, PI3 kinase pathway, and PLC pathway, transmitting positive signals like enhanced survival and growth. However, neurotrophin signaling involvement in AS is not well defined; studies have rarely been reported previously.

There are several important implications of this work. The identification of a group of small molecules with potential therapeutic efficacy for Ankylosing Spondylitis is an important observation. A total of 15 small molecules were identified have common sub-pathways with AS, including anisomycin (number of overlaps = 8), vorinostat (number of overlaps = 5), trichostatin A (number of overlaps = 5).

Anisomycin is an antibiotic produced by *Streptomyces griseolus* which inhibits protein synthesis [17]. Previous studies suggested that anisomycin can activate stress-activated protein kinases, MAP kinase and other signal transduction pathways. In our study (Figure 1), anisomycin was identified to have eight common sub-pathways (corresponding to five entire pathways) with AS, including B cell receptor signaling pathway, Fc epsilon RI signaling pathway and natural killer cell mediated cytotoxicity. It is mentioned as a potential psychiatric drug, as it may inhibit the consolidation of new context-specific long-term memories [18]. Vorinostat is a member of a larger class of compounds that inhibits histone deacetylases. It is the first drug approved for the treatment of cutaneous manifestations in patients with cutaneous T cell lymphoma. From Figure 1, we could easily find that pathways in cancer, cell cycle and Fc epsilon RI signaling pathway were perturbed by vorinostat. In clinical trials, vorinostat has shown significant anticancer activity against both hematologic and solid tumors at doses well tolerated by patients [19]. Trichostatin A inhibits the eukaryotic cell cycle during the beginning of the growth stage and has some potential as an anti-cancer drug [20]. In agreement with previous study, our study found that trichostatin A may perturb pathways of cell cycle, pathways in cancer and Fc gamma R-mediated phagocytosis. We are not aware of any reports of the evaluation or use of these compounds as systemic therapies for AS. Although these compounds are widely used in clinical practice, there is a lack of knowledge as to their effects on AS. Given the widespread investigation of these agents for the treatment of cancers, it is unlikely that they will be exploited as therapeutics by themselves, but may have promise for use in combination therapy. These observations warrant further study, and should generate hypotheses for laboratory, patient or population-based studies.

In conclusion, our study showed that neurotrophic signaling pathway and some immune-associated pathways may involve in the development of AS. Besides, our bioinformatics analysis revealed a total of 15 small molecules which may play a role in perturbing the development of AS. Our study proposed the use of an approach based on targeting sub-pathways to identify potential agents for AS. Further experimental verification is still needed to make sure that the use of these agents is effective without inhibiting important host defense mechanisms.

4. Experimental

4.1. Microarray Data

The transcription profile of GSE25101 was obtained from the National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database [21] which was based on the Illumina Inc. GPL6947 platform data (Illumina Human HT-12 whole-genome expression BeadChips). A total of

16 active AS patients, classified according to the New York criteria [22] and 16 gender-and age-matched controls were profiled using Illumina HT-12 Whole-Genome Expression BeadChips which carry cDNAs for 48000 genes and transcripts. Included patients had Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) scores > 4 and Bath Ankylosing Spondylitis Functional Index (BASFI) scores > 4 . We downloaded the raw data and the probe annotation files for further analysis.

4.2. Pathway Data

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals [23]. The PATHWAY database records networks of molecular interactions in the cells, and variants of them specific to particular organisms [24]. A total of 130 pathways, involving 2,287 genes, were collected from KEGG on June 30, 2011.

4.3. Small Molecules Data

The connectivity map (CMap) can be used to find connections among small molecules sharing a mechanism of action, chemicals and physiological processes, and diseases and drugs [25]. It is the first installment of a reference collection of gene-expression profiles from cultured human cells treated with bioactive small molecules, together with pattern-matching software to mine these data. The CMap dataset comprises genomic profiling data from 6,100 treatment-control pairs (instances) involving 1,309 bioactive molecules (perturbagens).

4.4. Differentially Expressed Genes Analysis

The classical t-test was used to identify the differentially expressed genes (DEGs). We used the Affy package in R [26] to preprocess the data of profile GSE25101. The raw expression datasets from all conditions were scaled using RMA method [27] with the default setting implemented in Bioconductor, and then the linear model was constructed. For each sample, the expression values of all probes for a given gene were reduced to a single value by taking the average expression value. To circumvent the multi-test problem which might induce too much false positive results, the BH method [28] was used to adjust the raw p . The p -value less than 0.1 was chosen as the cutoff criterion for DEGs.

4.5. Obtainment of Sub-Pathways by Parsing the KEGG pathway

SubpathwayMiner is an R-based software [26] package which facilitates sub-pathway identification of metabolic pathways by using pathway structure information. The sub-pathway is defined by an individual path from a start-point to an end-point in a pathway map. For a given KEGG pathway, the sub-pathways were obtained by searching all possible paths between start-points (membrane receptors or their ligands) and end-points (transcriptional factors or their immediate targets) in the adjacency matrix generated by node relationships [10].

The metabolic pathways downloaded from KEGG were converted to undirected graph with enzymes as nodes. Two nodes in an undirected graph are connected by an edge if there is a common compound in the enzymes corresponding reactions. To find the sub-pathways in which all enzymes have highly similar functions, we adopt the k-clique concept in social network analysis to define

sub-pathways based on distance similarity among enzymes [29]. In social network analysis, a k-clique in a graph is considered as a sub-graph where the distance between any two nodes is no greater than a parameter, k. The distance among all enzymes in mined sub-pathways decreases as the value of the parameter k reduces. Here, we set $k = 3$.

5. Conclusions

Our study proposes the use of an approach based on targeting sub-pathways to identify potential agents for AS. Candidate agents identified by our approach may provide the groundwork for a combination therapy approach for AS.

References

1. Schett, G. Bone formation versus bone resorption in Ankylosing Spondylitis. *Adv. Exp. Med. Biol.* **2009**, *649*, 114–121.
2. Assassi, S.; Reveille, J.D.; Arnett, F.C.; Weisman, M.H.; Ward, M.M.; Agarwal, S.K.; Gourh, P.; Bhula, J.; Sharif, R.; Sampat, K.; *et al.* Whole-blood gene expression profiling in Ankylosing Spondylitis shows upregulation of toll-like receptor 4 and 5. *J. Rheumatol.* **2011**, *38*, 87–98.
3. Reveille, J.D. Major histocompatibility genes and Ankylosing Spondylitis. *Best Pract. Res. Clin. Rheumatol.* **2006**, *20*, 601–609.
4. Burton, P.R.; Clayton, D.G.; Cardon, L.R.; Craddock, N.; Deloukas, P.; Duncanson, A.; Kwiatkowski, D.P.; McCarthy M.I.; Ouwehand, W.H.; Samani, N.J.; *et al.* Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat. Genet.* **2007**, *39*, 1329–1337.
5. Brown, M.A. Genetics and the pathogenesis of Ankylosing Spondylitis. *Curr. Opin. Rheumatol.* **2009**, *21*, 318–323.
6. Toussirot, E.; Wendling, D. Recent progress in Ankylosing Spondylitis treatment. *Expert. Opin. Pharmacother.* **2003**, *4*, 1–12.
7. Henes, J.C.; Horger, M.; Guenaydin, I.; Kanz, L.; Koetter, I. Mixed response to tocilizumab for Ankylosing Spondylitis. *Ann. Rheum. Dis.* **2010**, *69*, 2217–2218.
8. Rodriguez-Escalera, C.; Fernandez-Nebro, A. The use of rituximab to treat a patient with ankylosing spondylitis and hepatitis B. *Rheumatology (Oxford)* **2008**, *47*, 1732–1733.
9. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13.
10. Chen, X.; Xu, J.; Huang, B.; Li, J.; Wu, X.; Ma, L.; Jia, X.; Bian, X.; Tan, F.; Liu, L.; *et al.* A sub-pathway-based approach for identifying drug response principal network. *Bioinformatics* **2011**, *27*, 649–654.
11. Rihl, M.; Baeten, D.; Seta, N.; Gu, J.; Keyser, F.D.; Veys, E.; Kuipers, J.; Zeidler, H.; Yu, D. Technical validation of cDNA based microarray as screening technique to identify candidate genes in synovial tissue biopsy specimens from patients with spondyloarthritis. *Ann. Rheum. Dis.* **2004**, *63*, 498–507.
12. Rihl, M.; Kruihof, E.; Barthel, C.; Keyser, F.; Veys, E.M.; Zeidler, H.; Yu, D.T.; Kuipers, J.G.; Baeten, D. Involvement of neurotrophins and their receptors in spondyloarthritis synovitis: Relation to inflammation and response to treatment. *Ann. Rheum. Dis.* **2005**, *64*, 1542–1549.

13. Huang, E.J.; Reichardt, L.F. Neurotrophins: Roles in neuronal development and function. *Annu. Rev. Neurosci.* **2001**, *24*, 677–736.
14. Birling, M.C.; Price, J. Influence of growth factors on neuronal differentiation. *Curr. Opin. Cell Biol.* **1995**, *7*, 878–884.
15. Hempstead, B.L. Dissecting the diverse actions of pro- and mature neurotrophins. *Curr. Alzheimer Res.* **2006**, *3*, 19–24.
16. Reichardt, L.F. Neurotrophin-regulated signalling pathways. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2006**, *361*, 1545–1564.
17. Grollman, A.P. Inhibitors of protein biosynthesis. II. Mode of action of anisomycin. *J. Biol. Chem.* **1967**, *242*, 3226–3233.
18. Barrientos, R.M.; O'Reilly, R.C.; Rudy, J.W. Memory for context is impaired by injecting anisomycin into dorsal hippocampus following context exploration. *Behav. Brain Res.* **2002**, *134*, 299–306.
19. Marks, P.A.; Breslow, R. Dimethyl sulfoxide to vorinostat: Development of this histone deacetylase inhibitor as an anticancer drug. *Nat. Biotechnol.* **2007**, *25*, 84–90.
20. Drummond, D.C.; Noble, C.O.; Kirpotin, D.B.; Guo, Z.; Scott, G.K.; Benz, C.C. Clinical development of histone deacetylase inhibitors as anticancer agents. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 495–528.
21. GEO Homepage. Available online: <http://www.ncbi.nlm.nih.gov/geo/> (accessed on 10 October 2012).
22. Van Der Linden, S.; Valkenburg, H.A.; Cats, A. Evaluation of diagnostic criteria for Ankylosing Spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum.* **1984**, *27*, 361–368.
23. Kanehisa, M. The KEGG database. *Novartis Found. Symp.* **2002**, *247*, 91–101.
24. KEGG Homepage. Available online: <http://www.genome.jp/kegg/> (accessed on 10 October 2012).
25. Lamb, J.; Crawford, E.D.; Peck, D.; Modell, J.W.; Blat, I.C.; Wrobel, M.J.; Lerner, J.; Brunet, J.P.; Subramanian, A.; Ross, K.N.; *et al.* The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **2006**, *313*, 1929–1935.
26. *SubpathwayMiner-R-Based Software*; A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2011.
27. Irizarry, R.A.; Hobbs, B.; Collin, F.; Beazer-Barclay, Y.D.; Antonellis, K.J.; Scherf, U.; Speed, T.P. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **2003**, *4*, 249–264.
28. Benjamini, Y.H.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **1995**, *57*, 289–300.
29. Li, C.; Li, X.; Miao, Y.; Wang, Q.; Jiang, W.; Xu, C.; Li, J.; Han, J.; Zhang, F.; Gong, B.; *et al.* SubpathwayMiner: A software package for flexible identification of pathways. *Nucleic Acids Res.* **2009**, e131.

Sample Availability: Not available.