Supplementary table 5: Comparison of the precision efficiencies

S. no.	Tools	**Query set 1*:** ("RNA binding protein" OR "RNA binding proteins") AND (transcription initiation)		**Query set 2*:** ("alternative promoter" OR "alternative promoters") AND (mouse OR mice OR mus)		**Query set 3*:** Liver AND toxicity AND "cell death"		**Indicative precision Score (on a 1 to 10 scale)**
		Score for search I	**Equivalence score**	**Score for search II**	**Equivalence score**	**Score for search III**	**Equivalence score**	
1	**askMEDLINE**	28	4	46.7	1	93.3	5	3.3
2	**BioAsk**	26	3	115	7	117.1	7	5.7
3	**Brij.in**	36	7	86.7	4	26.7	1	4.0
4	**CiteXplore**	46	10	136.7	8	103	6	**8.0**
5	**EBIMed**	46	10	121.7	7	130	8	**8.3**
6	**eTBLAST**	28	4	-	-	-	-	3.0
7	**Google Scholar**	48	10	81.4	4	122.9	7	7.0
8	**HighWire Press**	48	10	105.7	6	123.6	7	7.7
9	**PubMed**	42	9	126.7	8	93	5	7.3
10	**PubMed Central**	-	-	-	-	65.7	3	3.0
11	**ReleMed**	32	5	162.9	10	169.2	10	**8.3**
12	**Scirus**	18	1	36.4	1	64.3	3	1.7
13	**Scopus****	42	9	143.3	9	150.7	9	**9.0**
14	**XplorMed**	38	7	116.7	7	110	6	6.6

*Please refer to supplementary notes 1 for a complete list of search terms used in the current study, along with the specific objectives.

**A trial version of the Scopus was used. This paid tool couldn't be used for the second round studies (supplementary notes 3) due to limited period of the trial access.

Please see supplementary notes 1 for details of the scoring system.

Notes:
1. PubMed interfaces (PubMed Assistant, PubMed Interact, PubFocus, GoPubMed, ClusterMed, BioIE and ALIBABA) were not included in the analysis as their primary output is similar to that of PubMed.

2. Though we began the initial studies with specific biological objectives (e.g., to list RNA binding proteins that interact with DNA in the promoter region, and gather reports available on role of the RNA binding proteins in transcription initiation), we suspected a danger of the effort being less useful for most of the readers who may not interested in the specific questions raised. Hence we shifted our focus on comparing the functionalities of the search engines in general, which is an urgent need of the scientific community today.

3. No single topic, or even 10 different topics, can represent the searches by all researchers and clinicians. Our objective of comparing the retrieval efficiency in general can be best met with 3 or more well defined query topics for each of the several broad subject areas. Choosing multiple query sets would avoid potential bias due to one particular set of query terms. Choosing a large number of broad areas (infectious diseases, neurological disorders, computational structural biology, system biology, algorithms, physiology, molecular biology etc) can be very good. But considering the volume of work required for each query, we restricted ourselves to cell and molecular biology, and then identified 3 query topics in an attempt to represent diversity within this broad subject. The expertise of the researchers involved in the work was then considered. It was necessary to be able and confident in judging the relevance of articles. KA has been involved in research mainly in the area of transcriptional regulation in mammalian male reproductive system. HH was involved in RNA binding protein studies and methods of RNA isolation.

4. Instead of fine tuning the query set to make best use of the search engine, we chose a simple set that is uniformly applicable across the search programs selected for citation retrieval efficiency comparisons. It is important to note that the results vary depending on the query set used.

5. Studies involving MeSH term efficiency, truncation features and subsets offered by PubMed, were independent of the other comparative studies and the above query topics were not the best in these cases. For example, MeSH seems to be biased towards

medical terms and hence we choose two medical topics, one of which is fact an example in tutorial of PubMed for MeSH usage, and one of general molecular biology significance.

6. When relevant, within the current approach, we have identified a specific objective and compared the efficiencies of the tools in that context.

7. All the query sets and purposes are now listed in supplementary notes 1.