
Supplementary information

**Learning on knowledge graph dynamics
provides an early warning of impactful
research**

In the format provided by the
authors and unedited

Learning on knowledge graph dynamics provides early warning of impactful research (Supplementary Material)

James W. Weis^{1,2} and Joseph M. Jacobson^{1,3}

¹MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

²Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

³MIT Center for Bits and Atoms, Massachusetts Institute of Technology, Cambridge, MA, USA

Table 1. The 42 life sciences journals from which research was identified and included in the biotechnology-graph analysis.

Angewandte Chemie	The Lancet
Blood	Nature Cell Biology
Cancer Cell	Nature Chemical Biology
Cancer Discovery	Nature Chemistry
Cancer Research	Nature Medicine
Cell	Nature Methods
Cell Host & Microbe	Nature
Cell Metabolism	Nature Biotechnology
Cell Stem Cell	The New England Journal of Medicine
Cell Chemical Biology	Neuron
The EMBO Journal	Nature Genetics
Genes & Development	Nature Immunology
Immunity	Nature Neuroscience
Journal of Neurology	Nature Structural & Molecular Biology
Journal of the American Chemical Society	PLOS Biology
JAMA	PLOS Genetics
Journal of Biological Chemistry	PLOS Pathogens
Journal of Cell Biology	Proceedings of the National Academy of Sciences
Journal of Clinical Investigation	Science Signaling
Journal of Experimental Medicine	Science Translational Medicine
Journal of Medicinal Chemistry	Science

Table 2. A sampling of recent publications predicted to be of future high-impact by DELPHI. A DELPHI model was trained on a restricted version of the biotechnology-focused dataset which only contained data from 2017 and earlier. Metrics were then extracted for articles in our dataset published in 2018 for 2018 and 2019, and the new DELPHI model was used to rank the publications by predicted probability of high-impact. Despite the fact that this analysis is limited to features extracted from the sub-graph of 42 journals, the identified publications including research topics spanning beyond biotechnology. Citation counts as of 2019 from within our 42 journal dataset, as well as the full academic graph, are shown. We note that this table is a sampling of publications identified by DELPHI to be of likely high future impact; not all papers identified by DELPHI as highly impactful appear on this table.

Title	Journal	Dataset Citations	Full-graph Citations
A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo	Nature Biotechnology	24	212
A high-energy-density lithium-oxygen battery based on a reversible four-electron conversion to lithium oxide	Science	13	60
A living biobank of breast cancer organoids captures disease heterogeneity	Cell	24	213
An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities	Nature Biotechnology	11	88
An in vivo model of functional and vascularized human brain organoids	Nature Biotechnology	18	155
Analysis of shared heritability in common disorders of the brain	Science	20	329
Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors	Nature Biotechnology	42	239
Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates	Nature	26	140
Coactivator condensation at super-enhancers links phase separation and gene control	Science	53	267
CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity	Science	24	208
De novo DNA synthesis using polymerase-nucleotide conjugates	Nature Biotechnology	4	49
Design of Single-Atom Co-N ₅ Catalytic Site: A Robust Electrocatalyst for CO ₂ Reduction with Nearly 100% CO Selectivity and Remarkable Stability	JACS	19	133
Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria	Nature Biotechnology	10	53
Enhancing the potential of enantioselective organocatalysis with light	Nature	55	97
Evolved Cas9 variants with broad PAM compatibility and high DNA specificity	Nature	36	367
Fluorine and Fluorinated Motifs in the Design and Application of Bioisosteres for Drug Design	Journal of Medicinal Chemistry	33	215
Genetic identification of brain cell types underlying schizophrenia	Nature genetics	16	108
Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma	NEJM	24	251
Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma	Science	30	199
Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017	The Lancet	28	423

Table 2 continued on the following page

Table 2 continued from the previous page

Title	Journal	Dataset Citations	Full-graph Citations
H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers	Nature Medicine	18	93
Health Care Spending in the United States and Other High-Income Countries	JAMA	29	347
Host expression of PD-L1 determines efficacy of PD-L1 pathway blockade-mediated tumor regression	Journal of Clinical Investigation	19	102
Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown	Cell	16	79
Humanized TREM2 mice reveal microglia-intrinsic and -extrinsic effects of R47H polymorphism	Journal of Experimental Medicine	16	64
Hypothalamic Circuits for Predation and Evasion	Neuron	15	48
Imaging dynamic and selective low-complexity domain interactions that control gene transcription	Science	27	145
Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs	PLoS pathogens	4	24
Integrating single-cell transcriptomic data across different conditions, technologies, and species	Nature Biotechnology	205	1040
Ionic liquids for oral insulin delivery	PNAS	7	42
Itaconate is an anti-inflammatory metabolite that activates Nrf2 via alkylation of KEAP1	Nature	34	179
Mapping the Mouse Cell Atlas by Microwell-Seq	Cell	66	279
Miniaturized neural system for chronic, local intracerebral drug delivery	Science Translational Medicine	8	28
Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain	Cell	29	168
mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding	Cell	11	49
NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control	Cell	21	177
Planning chemical syntheses with deep neural networks and symbolic AI	Nature	22	274
Regulation of age-associated B cells by IRF5 in systemic autoimmunity	Nature Immunology	5	24
RNA velocity of single cells	Nature	37	219
Semiconducting Polymer Nanoenzymes with Photothermal Activity for Enhanced Cancer Therapy	Angewandte Chemie	16	82
Shared and distinct transcriptomic cell types across neocortical areas	Nature	26	180
Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars	Nature Biotechnology	21	103
Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain	Nature Biotechnology	22	135
Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment	Cell	42	204
Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease	Science	18	175
STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma	Cancer Discovery	14	133

Table 2 continued on the following page

Table 2 continued from the previous page

Title	Journal	Dataset Citations	Full-graph Citations
Tau protein liquid-liquid phase separation can initiate tau aggregation	The EMBO Journal	20	117
The human transcription factors	Cell	24	289
The long tail of oncogenic drivers in prostate cancer	Nature genetics	24	142
Thermal proximity coaggregation for system-wide profiling of protein complex dynamics in cells	Science	7	29
Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors	Science	21	127
Unexpected reversal of C3 versus C4 grass response to elevated CO ₂ during a 20-year field experiment	Science	8	73
VPS13A and VPS13C are lipid transport proteins differentially localized at ER contact sites	Journal of Cell Biology	10	61