nature REVIEWS DRUG DISCOVERY

SUPPLEMENTARY INFORMATION

In format as provided by the authors

Mapping biologically active chemical space to accelerate drug discovery

G. Sitta Sittampalam, Dobrila D. Rudnicki, Danilo A. Tagle, Anton Simeonov and Christopher P. Austin

https://doi.org/10.1038/d41573-018-00007-2

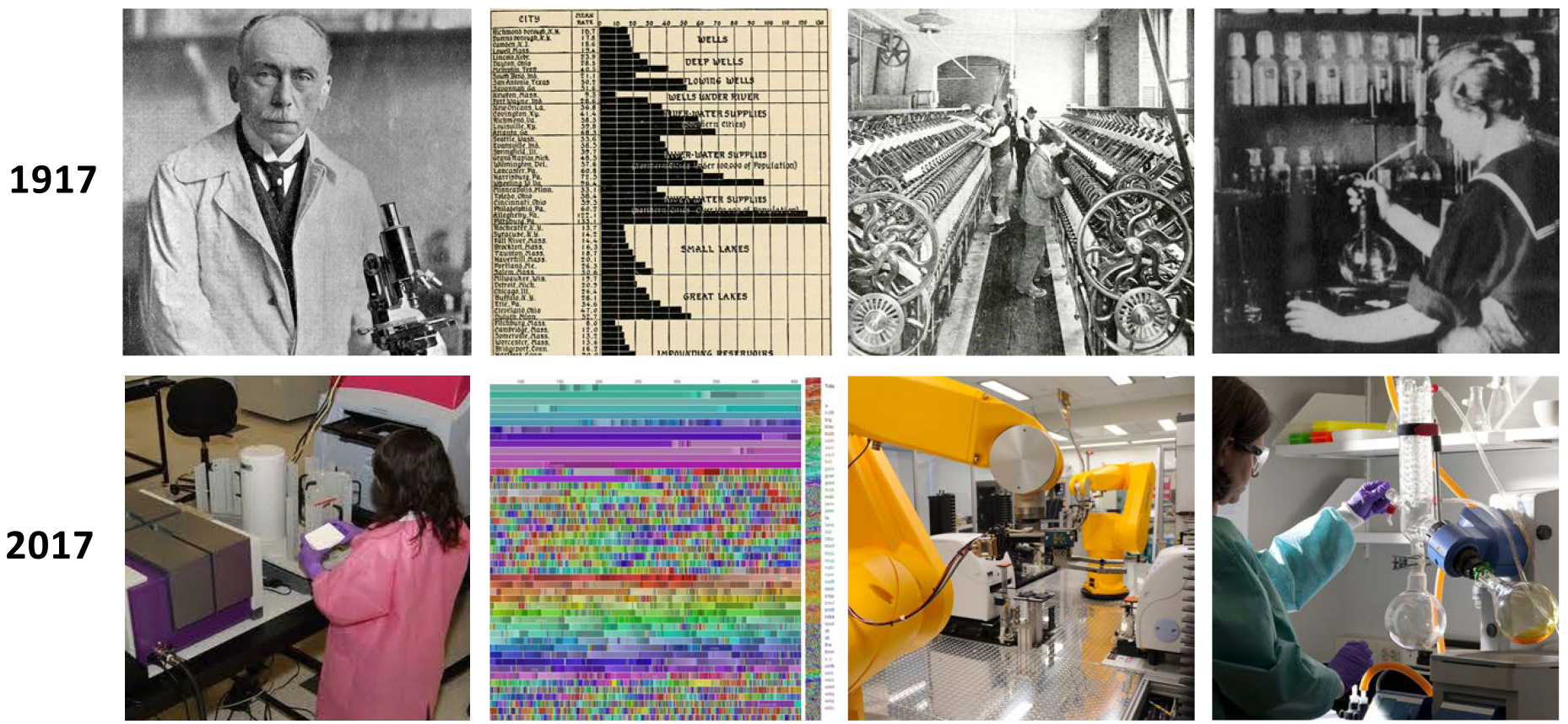

Figure by Kyle Brimacombe, NIH/NCATS

Supplementary Figure S1 | **Illustration of the progression in key fields for biomedical research in the past century.**

1917 Biology image reproduced from https://wellcomecollection.org/works/hxd7rtq9?query=Jules+Jean+Baptiste+Vincent+Bordet, under CC-BY 4.0 license (https://creativecommons.org/licenses/by/4.0/). 1917 Informatics image reproduced from 'Preventive medicine and hygiene' by Rosenau, M. J. (Milton Joseph), 1869–1946; Whipple, George Chandler, 1866–1924; Trask, John W. (John William), b. 1877; Salmon, Thomas William, published by New York and London D. Appleton and Company 1917. 1917 Engineering image reproduced from 'School of Industrial Art of the Pennsylvania Museum; Circular of the Philadelphia Textile School, Thirty-Fourth Season 1917–1918. 1917 Chemistry image adapted from Smithsonian Institution Archives. Image # SIA2010-0474. 2017 Biology image reproduced from https://www.flickr.com/photos/64860478@N05/5965218450/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/). 2017 Informatics image reproduced from https://commons.wikimedia.org/wiki/File:Viegas-UserActivityonWikipedia.gif, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/). 2017 Engineering image reproduced from https://commons.wikimedia.org/wiki/File:Chemical_Genomics_Robot.jpg/ https://www.flickr.com/photos/64860478@N05/5964663403/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/). 2017 Chemistry image reproduced from https://www.flickr.com/photos/nationaleyeinstitute/9955408003/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/).

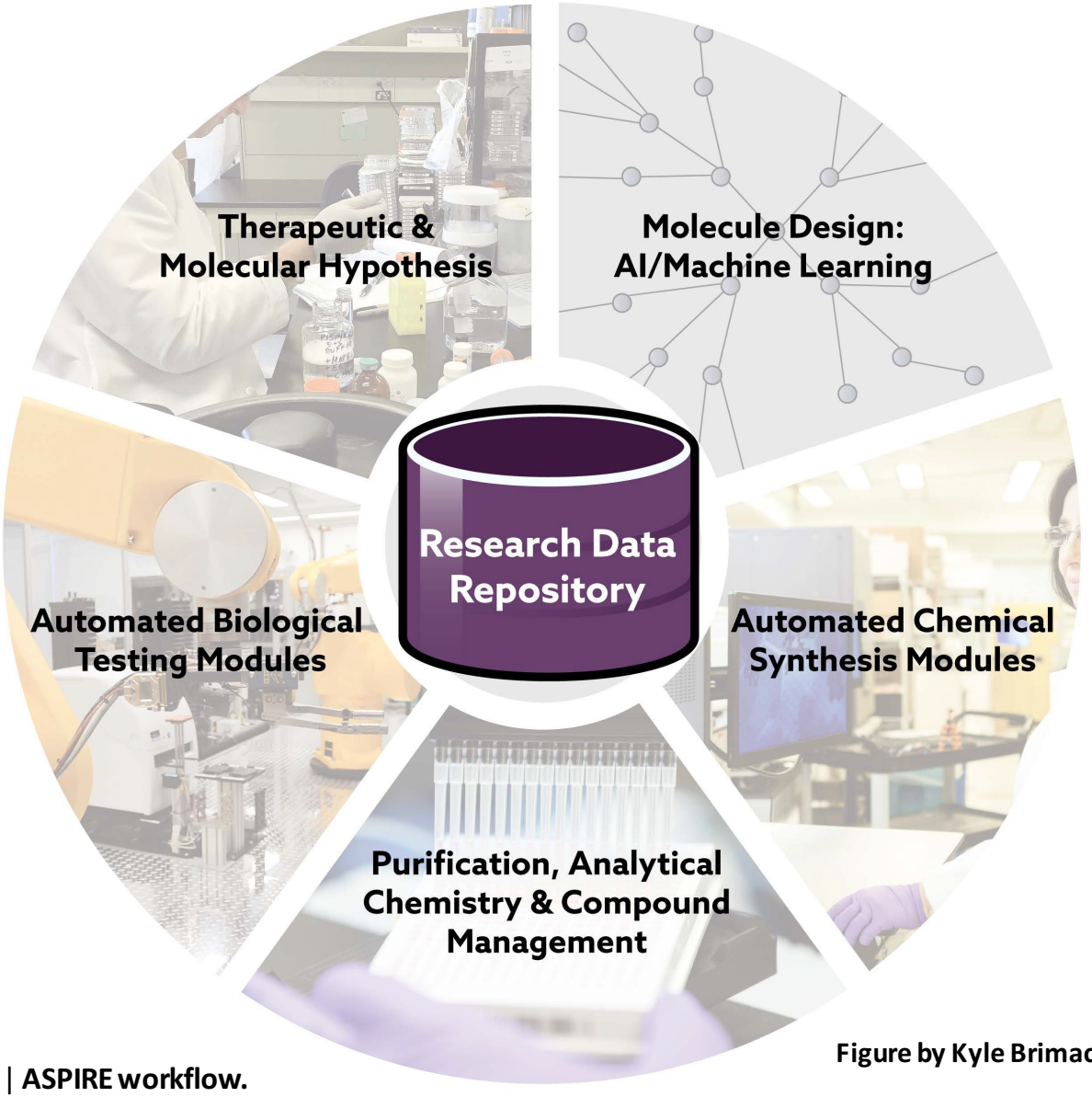

Figure by Kyle Brimacombe, NIH/NCATS

Supplementary Figure S2 | **ASPIRE workflow.**

Therapeutic & Molecular Hypothesis image adapted from https://www.flickr.com/photos/64860478@N05/26442537689/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/). Molecule Design: AI/Machine Learning image adapted from https://commons.wikimedia.org/wiki/File:Network_Tree_diagram.svg, under the Creative Commons CC0 1.0 Universal Public Domain Dedication (https://creativecommons.org/publicdomain/zero/1.0/). Automated Chemical Synthesis modules image adapted from https://www.flickr.com/photos/64860478@N05/38709125891/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/). Purification, Analytical Chemistry & Compound Management image adapted from https://www.flickr.com/photos/64860478@N05/38709125891/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/). Automated Biological Testing Modules image adapted from https://commons.wikimedia.org/wiki/File:Chemical_Genomics_Robot.jpg/https://www.flickr.com/photos/64860478@N05/5964663403/, under CC-BY 2.0 generic licence (https://creativecommons.org/licenses/by/2.0/).

Supplementary Table 1 | Automated Chemical Synthesis Workshop, 19–20 October 2017

Natcher Conference Center • National Institutes of Health • Bethesda, Maryland

External thought leader participation		NIH workshop organization	
Academic research	**Pharmaceutical/industry Research**	**NCATS**	**Role**
Carnegie Mellon University Joshua Kangas, Ph.D., Research Scientist, School of Computer Science	**HitGen Ltd** Barry A. Morgan, Ph.D., Chief Scientific Officer	Christopher P. Austin, M.D., Director	Workshop Sponsor, Session Chair, Speaker
Georgia Inst. Of Technology Jeffrey Skolnick, Ph.D., Professor, , Mary and Maisie Gibson Chair & GRA Eminent Scholar in Computational Systems Biology	**Eli Lilly & Co. Discovery Chemistry Research & Technologies** Alan D. Palkowitz, Ph.D., Vice President Alex Godfrey, Ph.D., Chemistry Automation *(Retired)* Scott Sheehan, Ph.D., Chief Scientific Officer	Pamela McInnes,D.D.S., Deputy Director	Workshop Sponsor
		Anton Simeonov, Ph.D., Scientific Director	Workshop Sponsor, Speaker, Session Chair
		Danilo Tagle, Ph.D., Director, Special Initiatives	Workshop Sponsor
Indiana University Nicola L. B. Pohl, Ph.D., Professor and Joan & Marvin Carmack Chair	**IBM Research** Akihiro Kishimoto, Ph.D. Research Staff Member	Dobrila D. Rudnicki, Ph.D., Special Initiatives, Program Director	Overall Workshop organization and management
Max Planck Institute for Colloids and Interface, Max Planck Society Peter H. Seeberger, Ph.D., Director Kerry Gilmore Ph.D. Group Leader, Biomolecular Systems	**Merck** Timothy Cernak, Ph.D., Associate Principal Scientist, Discovery Chemistry – Automation and Capabilities Enhancement Spencer Dreher, Ph.D., Principal Scientist, Chemistry, Catalysis and Automation	G. Sitta Sittampalam, Ph.D., Senior Advisor to the Director	Senior Advisor, Conference Organization, Moderator, Panelist
		Sam Michael, Ph.D., Chief Information Officer, Head of Automation	Consultant, Moderator, Session Chair
MIT Stephen L. Buchwald, Ph.D., Camille Dreyfus Professor of Chemistry, Associate Head, Department of Chemistry Klavs Jensen, Ph.D., Warren K. Lewis Professor of Chemical Engineering, Professor of Materials Science and Engineering Connor W. Coley, Graduate Student	**MilliporeSigma** Sarah Trice, Ph.D., Head, Commercial Development Daniel Boesch, Ph.D., Head, Chemical Synthesis **EMD Serano** Mireille Krier, Group Leader	Group Leaders: Juan Marugan, Ph.D. Chris LeClair, PhD.	Consultant, Panel Discussion Consultant, Panel Discussion
		Senior Medicinal Chemists: Samarjit Patnaik, Ph.D. Jason Rohde, Ph.D. Damien Duveau, Ph.D.	Consultant, Panel discussions Moderator, Panel discussions Consultant, Panel discussions
Stanford University Paul Wender, Ph.D., Bergstrom Professor of Chemistry	**Publishers**	**NIGMS**	
The Scripps Research Institute Phil S. Baran, Ph.D., Darlene Shiley Chair in Chemistry, Professor, Department of Chemistry	**ACS Publication Division** Darla Henderson, Ph.D., Assistant Director, Open Access Programs	Jon R. Lorsch, Ph.D., Director	Co-organizer, Session Chair
University of Glasgow Leroy "Lee" Cronin, Ph.D., Regius Chair of Chemistry, School of Chemistry	**Beilstein Institute** Martin Hicks, Ph.D., Board of Management	Pamela A. Marino, Ph.D., Chief, Division of Pharmacology, Physiology, and Biological Chemistry	Co-organizer, Speaker, Session Chair
University of Illinois Martin D. Burke, Ph.D., Professor, Department of Chemistry	**Science** Jake Yeston, Ph.D., Deputy Editor for Phys. Sciences Research	Robert "Bob" Lees, Ph.D., Program Director	Co-organizer
		DARPA	
Ulsan National Institute of Science & Technology , S. Korea Bartosz A. Grzybowski, Ph.D., Distinguished Professor	Caroline Trupp Gil Director, Federal Relations, External Affairs and Communications	Anne Fischer, Ph.D., Program Manager, Defense Sciences Office	Co-organizer, Speaker, Panelist

Supplementary Table 2 | Major gaps, key stakeholders and possible solutions identified at the Automated Chemical Synthesis Workshop

Gaps identified	Stakeholders	Possible solutions
Data standards for reports and chemical structure representation for automated data capture.	Software developers, Data standards scientists, Machine learning and AI scientists, medicinal/synthetic chemists, publishers and professional societies.	Engage medicinal chemistry and AI community through their professional societies to develop data standards. Open source data bases through global engagement.
Large data sets for machine learning predictions (big chemistry data sets). Open data repositories and data bases. Positive and negative data of synthetic reaction conditions to improve reproducibility and reaction predictions.	Machine learning and AI scientists, software/data repository developers, medicinal/synthetic chemists. Analytical chemists.	Develop inexpensive software tools (e.g. flexible and more versatile electronic laboratory notebooks (eLNs) for synthetic chemists to capture both positive and negative data simultaneously. Develop common definitions for positive versus negative data when evaluating synthetic reactions.
AI and deep learning algorithms to predict and optimize known and new reaction conditions, scope and limitations of reactions, and novel chemical libraries.	Medicinal/synthetic chemists, software developers, machine learning and AI scientists, analytical chemists, publishers and funders	New, flexible funding mechanisms for collaborations with academic and industrial scientists to develop reaction prediction algorithms using AI and machine learning concepts.
Automated sample handling of compounds and reagents with various physicochemical properties and stability.	Automation engineers, material scientists, medicinal/synthetic chemists, compound management professionals.	Establish a consortium of instrumentation vendors, material scientists, analytical chemists and medicinal chemists to develop new technologies for sample handling. (e.g. HTS equipment development from 1990s)
Technology for seamless integration of automated synthesis, purification, compound management and analytical characterization, and bioassays.	Automation engineers, instrument makers, software developers, analytical chemists, biologists, pharmacologists, compound management professionals, HTS scientists.	New, flexible funding mechanisms for collaborations with academic and industrial scientists to develop automated chemical synthesis, and biological annotation, such as the ASPIRE concept.
Need to address traditional medicinal chemistry culture, data and technology to enable integration of automated chemistry and biology. User-friendly automated chemistry modules and control software.	Medicinal/synthetic chemists, funders, publishers, professional societies, all drug discovery scientists, patient advocacy groups.	Work with chemical and automation/instrumentation industry to develop user-friendly synthetic automated chemistry modules with expanded sample handling and analytical capabilities. Symposium sessions in professional society meetings on automated chemistry and biology solutions.