We have demonstrated the discovery of physical laws, from scratch, directly from experimentally captured data with the use of a computational search. We used the presented approach to detect nonlinear energy conservation laws, Newtonian force laws, geometric invariants, and system manifolds in various synthetic and physically implemented systems without prior knowledge about physics, kinematics, or geometry. The concise analytical expressions that we found are amenable to human interpretation and help to reveal the physics underlying the observed phenomenon. Many applications exist for this approach, in fields ranging from systems biology to cosmology, where theoretical gaps exist despite abundance in data.

Might this process diminish the role of future scientists? Quite the contrary: Scientists may use processes such as this to help focus on interesting phenomena more rapidly and to interpret their meaning.

References and Notes

1. P. W. Anderson, *Science* **177**, 393 (1972).
2. E. Noether, *Nachr. d. König Gesellsch. d. Wiss. zu Göttingen, Math-Phys. Klasse* 235 (1918).
3. J. Hanc, S. Tuleja, M. Hancova, *Am. J. Phys.* **72**, 428 (2004).
4. D. Clery, D. Voss, *Science* **308**, 809 (2005).
5. A. Szalay, J. Gray, *Nature* **440**, 413 (2006).
6. R. E. Valdés-Pérez, *Commun. Assoc. Comput. Mach.* **42**, 37 (1999).
7. R. D. King *et al.*, *Nature* **427**, 247 (2004).
8. P. Langley, *Cogn. Sci.* **5**, 31 (1981).
9. R. M. Jones, P. Langley, *Comput. Intell.* **21**, 480 (2005).
10. J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. (MIT Press, Cambridge, MA, 1992).
11. S. Forrest, *Science* **261**, 872 (1993).
12. J. Duffy, J. Engle-Warnick, *Evolutionary Computation in Economics and Finance* **100**, 61 (2002).
13. F. Cyril, B. Alberto, in *2007 IEEE Congress on Evolutionary Computation*, S. Dipti, W. Lipo, Eds. (IEEE Press, Singapore, 2007), pp. 23–30.
14. B. Elena, B. Andrei, L. Henri, in *Seventh International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC '05)* (IEEE Press, 2005), pp. 321–324.
15. J. Bongard, H. Lipson, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9943 (2007).
16. S. Nee, N. Colegrave, S. A. West, A. Grafen, *Science* **309**, 1236 (2005).
17. P. Jäckel, T. Mullin, *Proc. R. Soc. London Ser. A* **454**, 3257 (1998).
18. T. Shinbrot, C. Grebogi, J. Wisdom, J. A. Yorke, *Am. J. Phys.* **60**, 491 (1992).
19. Y. Liang, B. Feeny, *Nonlinear Dyn.* **52**, 181 (2008).
20. M. Mor, A. Wolf, O. Gottlieb, in *Proceedings of the 21st ASME Biennial Conference on Mechanical Vibration and Noise* (ASME Press, Las Vegas, NV, 2007), pp. 1–8.
21. P. Gregory, R. Denis, F. Cyril, in *Evolution Artificielle, 6th International Conference*, vol. 2936, L. Pierre, C. Pierre, F. Cyril, L. Evelyne, S. Marc, Eds. (Springer, Marseilles, France, 2003), pp. 267–277.
22. E. D. De Jong, J. B. Pollack, in *Genetic Programming and Evolvable Machines*, vol. 4 (Springer, Berlin, 2003), pp. 211–233.
23. S. H. Strogatz, *Nature* **410**, 268 (2001).
24. P. A. Marquet, *Nature* **418**, 723 (2002).

25. This research was supported in part by Integrative Graduate Education and Research Traineeship program in nonlinear systems, a U.S. NSF graduate research fellowship, and NSF Creative-IT grant 0757478 and CAREER grant 0547376. We thank M. Kurman for editorial consultation and substantive editing of the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/324/5923/81/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Tables S1 to S3
References
Movie S1
Data Sets S1 to S15

15 September 2008; accepted 19 February 2009
10.1126/science.1165893

The Automation of Science

Ross D. King,[1]* Jem Rowland,[1] Stephen G. Oliver,[2] Michael Young,[3] Wayne Aubrey,[1] Emma Byrne,[1] Maria Liakata,[1] Magdalena Markham,[1] Pınar Pir,[2] Larisa N. Soldatova,[1] Andrew Sparkes,[1] Kenneth E. Whelan,[1] Amanda Clare[1]

The basis of science is the hypothetico-deductive method and the recording of experiments in sufficient detail to enable reproducibility. We report the development of Robot Scientist "Adam," which advances the automation of both. Adam has autonomously generated functional genomics hypotheses about the yeast *Saccharomyces cerevisiae* and experimentally tested these hypotheses by using laboratory automation. We have confirmed Adam's conclusions through manual experiments. To describe Adam's research, we have developed an ontology and logical language. The resulting formalization involves over 10,000 different research units in a nested treelike structure, 10 levels deep, that relates the 6.6 million biomass measurements to their logical description. This formalization describes how a machine contributed to scientific knowledge.

Computers are playing an ever-greater role in the scientific process (*1*). Their use to control the execution of experiments contributes to a vast expansion in the production of scientific data (*2*). This growth in scientific data, in turn, requires the increased use of computers for analysis and modeling. The use of computers is also changing the way that science is described and reported. Scientific knowledge is best expressed in formal logical languages (*3*). Only formal languages provide sufficient semantic clarity to ensure reproducibility and the free exchange of scientific knowledge. Despite the advantages of logic, most scientific knowledge is expressed only in natural languages. This is now changing through developments such as the Semantic Web (*4*) and ontologies (*5*).

A natural extension of the trend to ever-greater computer involvement in science is the concept of a robot scientist (*6*). This is a physically implemented laboratory automation system that exploits techniques from the field of artificial intelligence (*7*–*9*) to execute cycles of scientific experimentation. A robot scientist automatically originates hypotheses to explain observations, devises experiments to test these hypotheses, physically runs the experiments by using laboratory robotics, interprets the results, and then repeats the cycle.

High-throughput laboratory automation is transforming biology and revealing vast amounts of new scientific knowledge (*10*). Nevertheless, existing high-throughput methods are currently inadequate for areas such as systems biology. This is because, even though very large numbers of experiments can be executed, each individual experiment cannot be designed to test a hypothesis about a model. Robot scientists have the potential to overcome this fundamental limitation.

The complexity of biological systems necessitates the recording of experimental metadata in as much detail as possible. Acquiring these metadata has often proved problematic. With robot scientists, comprehensive metadata are produced as a natural by-product of the way they work. Because the experiments are conceived and executed automatically by computer, it is possible to completely capture and digitally curate all aspects of the scientific process (*11*, *12*).

To demonstrate that the robot scientist methodology can be both automated and be made effective enough to contribute to scientific knowledge, we have developed Robot Scientist "Adam" (*13*) (Fig. 1). Adam's hardware is fully automated such that it only requires a technician to periodically add laboratory consumables and to remove waste. It is designed to automate the high-throughput execution of individually designed microbial batch growth experiments in microtiter plates (*14*). Adam measures growth curves (phenotypes) of selected microbial strains (genotypes) growing in defined media (environments). Growth of cell cultures can be easily measured in high-throughput, and growth curves are sensitive to changes in genotype and environment.

We applied Adam to the identification of genes encoding orphan enzymes in *Saccharomyces cerevisiae*: enzymes catalyzing biochemical reactions thought to occur in yeast, but for which the encoding gene(s) are not known (*15*). To set up Adam for this application required (i) a comprehensive logical model encoding knowledge of *S. cerevisiae* metabolism [~1200 open

[1]Department of Computer Science, Aberystwyth University, SY23 3DB, UK. [2]Cambridge Systems Biology Centre, Department of Biochemistry, University of Cambridge, Sanger Building, 80 Tennis Court Road, Cambridge CB2 1GA, UK. [3]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, SY23 3DD, UK.

*To whom correspondence should be addressed. E-mail: rdk@aber.ac.uk

reading frames (ORFs), ~800 metabolites] (*15*), expressed in the logic programming language Prolog; (ii) a general bioinformatic database of genes and proteins involved in metabolism; (iii) software to abduce hypotheses about the genes encoding the orphan enzymes, done by using a combination of standard bioinformatic software and databases; (iv) software to deduce experiments that test the observational consequences of hypotheses (based on the model); (v) software to plan and design the experiments, which are based on the use of deletion mutants and the addition of selected metabolites to a defined growth medium; (vi) laboratory automation software to physically execute the experimental plan and to record the data and metadata in a relational database; (vii) software to analyze the data and metadata (generate growth curves and extract parameters); and (viii) software to relate the analyzed data to the hypotheses; for example, statistical methods are required to decide on significance. Once this infrastructure is in place, no human intellectual intervention is necessary to execute cycles of simple hypothesis-led experimentation. [For more details of the software, and its application to a related functional genomics problem, see (*16*) and figs. S1 and S2].

Adam formulated and tested 20 hypotheses concerning genes encoding 13 orphan enzymes (*16*) (Table 1). The weight of the experimental evidence for the hypotheses varied (based on observations of differential growth), but 12 hypotheses with no previous evidence were confirmed with $P < 0.05$ for the null hypothesis.

Because Adam's experimental evidence for its conclusions is indirect, we tested Adam's conclusions with more direct experimental methods. The enzyme 2-aminoadipate:2-oxoglutarate aminotransferase (2A2OA) catalyzes a reaction in the lysine biosynthetic pathways of fungi. Adam hypothesized that three genes (YER152C, YJL060W, and YGL202W) encode this enzyme and observed results consistent with all three hypotheses (Table 1). To test Adam's conclusions, we purified the protein products of these genes and used them in in vitro enzyme assays, which confirmed Adam's conclusions [supporting online material (SOM)] (Fig. 2).

To further test Adam's conclusions, we examined the scientific literature on the 20 genes investigated (Table 1) (*16*). This revealed the existence of strong empirical evidence for the correctness of six of the hypotheses; that is, the enzymes were not actually orphans (Table 1). The reason that Adam considered them to be orphans was due to the use of an incomplete bioinformatic database. These six genes therefore constitute a positive control for Adam's methodology. A possible error was also revealed (Table 1) (SOM).

To better understand the reasons why the identity of the genes encoding these enzymes has remained obscure for so long, we investigated their comparative genomics in detail (*16*). The likely explanation is a combination of three complicating factors: gene duplications with retention of overlapping function, enzymes that catalyze more than one related reaction, and existing functional annotations. Adam's systematic bioinformatic and quantitative phenotypic analyzes were required to unravel this web of functionality.

Use of a robot scientist enables all aspects of a scientific investigation to be formalized in logic. For the core organization of this formalization, we used the ontology of scientific experiments: EXPO (*11*, *12*). This ontology formalizes generic knowledge about experiments. For Adam, we developed LABORS, a customized version of EXPO, expressed in the description logic language OWL-DL (*17*). Application of LABORS produces experimental descriptions in the logic-

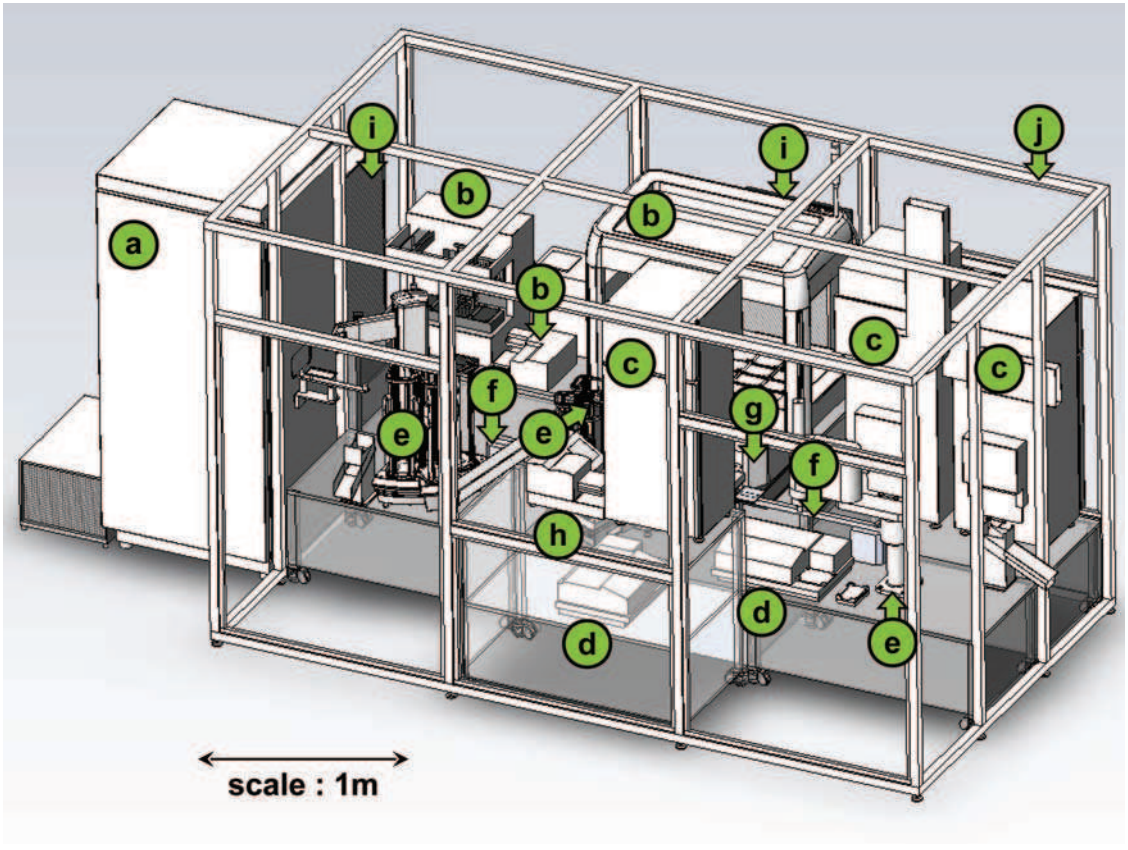

Fig. 1. The Robot Scientist Adam. The advances that distinguish Adam from other complex laboratory systems are the individual design of the experiments to test hypotheses and the utilization of complex internal cycles. Adam's basic operations are selection of specified yeast strains from a library held in a freezer, inoculation of these strains into microtiter plate wells containing rich medium, measurement of growth curves on rich medium, harvesting of a defined quantity of cells from each well, inoculation of these cells into wells containing defined media (minimal synthetic dextrose medium plus up to four added metabolites from a choice of six), and measurement of growth curves on the specified media. To achieve this functionality, Adam has the following components: a, an automated −20°C freezer; b, three liquid handlers (one of which can separately control 96 fluid channels simultaneously); c, three automated +30°C incubators; d, two automated plate readers; e, three robot arms; f, two automated plate slides; g, an automated plate centrifuge; h, an automated plate washer; i, two high-efficiency particulate air filters; and j, a rigid transparent plastic enclosure. There are also two bar code readers, seven cameras, 20 environment sensors, and four personal computers, as well as the software. Adam is capable of designing and initiating over a thousand new strain and defined-growth-medium experiments each day (from a selection of thousands of yeast strains), with each experiment lasting up to 5 days. The design enables measurement of OD_{595nm} for each experiment at least once every 30 min (more often if running at less than full capacity), allowing accurate growth curves to be recorded (typically we take over a hundred measurements a day per well), plus associated metadata. See the supporting online material for pictures and a video of Adam in action.

programming language Datalog (*18*). In the course of its investigations, Adam observed 6,657,024 optical density (OD_{595nm}) measurements (forming 26,495 growth curves). These data are held in a MySQL relational database. Use of LABORS resulted in a formalization of the scientific argumentation involving over 10,000 different research units (segments of experimental research). This has a nested treelike structure, 10 levels deep, that logically connects the experimental observations to the experimental metadata. (Fig. 3). This structure resembles the trace of a computer program and takes up 366 Mbytes (*16*). Making such experimental structures explicit renders scientific research more comprehensible, reproducible, and reusable. This paper may be considered as simply the human-friendly summary of the formalization.

Table 1. The orphan enzymes and Adam's hypotheses. The hypothesized genes are those which Adam abduced encoded an orphan enzyme. Prob. is Adam's Monte Carlo estimate of the probability of obtaining the observed discrimination accuracy or better with a random labeling of replicates. The discrimination is between the differences in growth curves observed with the addition of specified metabolites to the wild type and the deletant. Acc. is the highest accuracy for a metabolite species in discriminating between the growth curves observed with the addition of specified metabolites to the wild type and the deletant. No. is the number of metabolites tested. Existing annotation is the summary from the *Saccharomyces* Genome Database of the annotation of the ORF. Dry is the summary of whether the annotated function is the same as predicted by Adam. If a gene already has an associated function, we do not consider this to be contradictory to Adam's conclusions unless this function is capable of explaining the observed growth phenotype, for example, *BCY1*. ida indicates inferred from direct assay and iss, inferred from sequence or structural similarity (*5*). Wet is the result of our manual enzyme assays. See (*16*) for details.

Orphan enzyme	Hypothesized gene	Prob.	Acc.	No.	Existing annotation	Dry	Wet
Glucosamine-6-phosphate deaminase (3.5.99.6)	YHR163W (SOL3)	$<10^{-4}$	97	8	6-Phosphogluconolactonase, ida	–	–
Glutaminase (3.5.1.2)	YIL033C (BCY1)	$<10^{-4}$	92	11	Cyclic adenosine 3′,5′-monophosphate (cAMP)–dependent protein kinase inhibitor, ida	✗	–
L-Threonine 3-dehydrogenase (1.1.1.103)	YDL168W (SFA1)	$<10^{-4}$	83	6	Alcohol dehydrogenase, ida	–	–
Purine-nucleoside phosphorylase (2.4.2.1)	YLR209C (PNP1)	$<10^{-4}$	82	11	Purine-nucleoside phosphorylase, ida	✓	–
2-Aminoadipate transaminase (2.6.1.39)	YGL202W (ARO8)	$<10^{-4}$	80	3	Aromatic–amino acid transaminase, ida	✓	✓
5,10-Methenyltetrahydrofolate synthetase (6.3.3.2)	YER183C (FAU1)	$<10^{-4}$	80	4	5,10 Formyltetrahydrofolate cyclo-ligase, ida	✓	–
Glucosamine-6-phosphate deaminase (3.5.99.6)	YNR034W (SOL1)	$<10^{-4}$	79	2	Possible role in tRNA export	–	–
Pyridoxal kinase (2.7.1.35)	YPR121W (THI22)	$<10^{-4}$	78	1	Phosphomethylpyrimidine kinase, iss	–	–
Mannitol-1-phosphate 5-dehydrogenase (1.1.1.17)	YNR073C	$<10^{-4}$	78	6	Putative mannitol dehydrogenase, iss	–	–
1-Acylglycerol-3-phosphate O-acyltransferase (2.3.1.51)	YDL052C (SLC1)	0.0001	80	6	1-Acylglycerol-3-phosphate O-acyltransferase ida	✓	–
Glucosamine-6-phosphate deaminase (3.5.99.6)	YGR248W (SOL4)	0.0002	78	2	6-Phosphogluconolactonase, ida	–	–
Maleylacetoacetate isomerase (5.2.1.2)	YLL060C (GTT2)	0.0003	76	3	Glutathione S-transferase, ida	–	–
Serine O-acetyltransferase (2.3.1.30)	YJL218W	0.0005	78	2	Unknown function	–	–
L-Threonine 3-dehydrogenase (1.1.1.103)	YLR070C (XYL2)	0.0052	75	6	Xylitol dehydrogenase, ida	–	–
2-Aminoadipate transaminase (2.6.1.39)	YJL060W (BNA3)	0.0084	73	3	Kynurenine aminotransferase, ida	–	✓
Pyridoxal kinase (2.7.1.35)	YNR027W	0.0259	76	2	Involved in bud-site selection, iss	–	–
Polyamine oxidase (1.5.3.11)	YMR020W (FMS1)	0.0289	78	4	Polyamine oxidase, ida	✓	–
2-Aminoadipate transaminase (2.6.1.39)	YER152C	0.0332	74	3	Uncharacterized	–	✓
L-Aspartate oxidase (1.4.3.16)	YJL045W	0.1300	72	1	Succinate dehydrogenase isozyme, iss	–	–
Purine-nucleoside phosphorylase (2.4.2.1)	YLR017W (MEU1)	0.1421	72	6	Methylthioadenosine phosphorylase, ida	✓	–

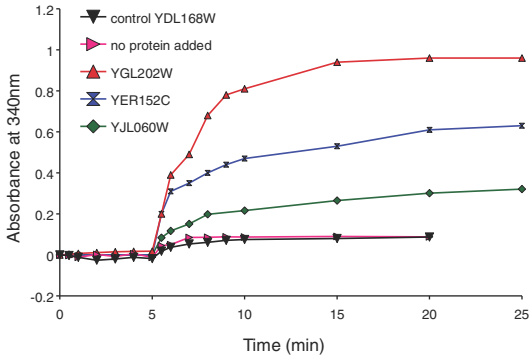

Fig. 2. Assay results for 2A2OA activity. The proteins encoded by YGL202W, YJL060W, YER152C, and YDL168W were expressed from OpenBiosystems (www.openbiosystems.com) yeast ORF clones and purified. Activity was tested in an assay of NADPH (reduced form of nicotinamide adenine dinucleotide phosphate) production based on (*22*). L-α-aminoadipic acid and 2-oxoglutarate were provided as substrates and pyridoxal phosphate as cofactor. Glutamate production was assayed by using commercially available yeast glutamate dehydrogenase, which uses NADP as cofactor and deaminates glutamate, producing ammonia and NADPH and regenerating 2-oxoglutarate (*16*). Also consistent with 2A2OA activity is experimental evidence indicating a higher activity with L-α-aminoadipic acid over either alanine or aspartate (*16*).

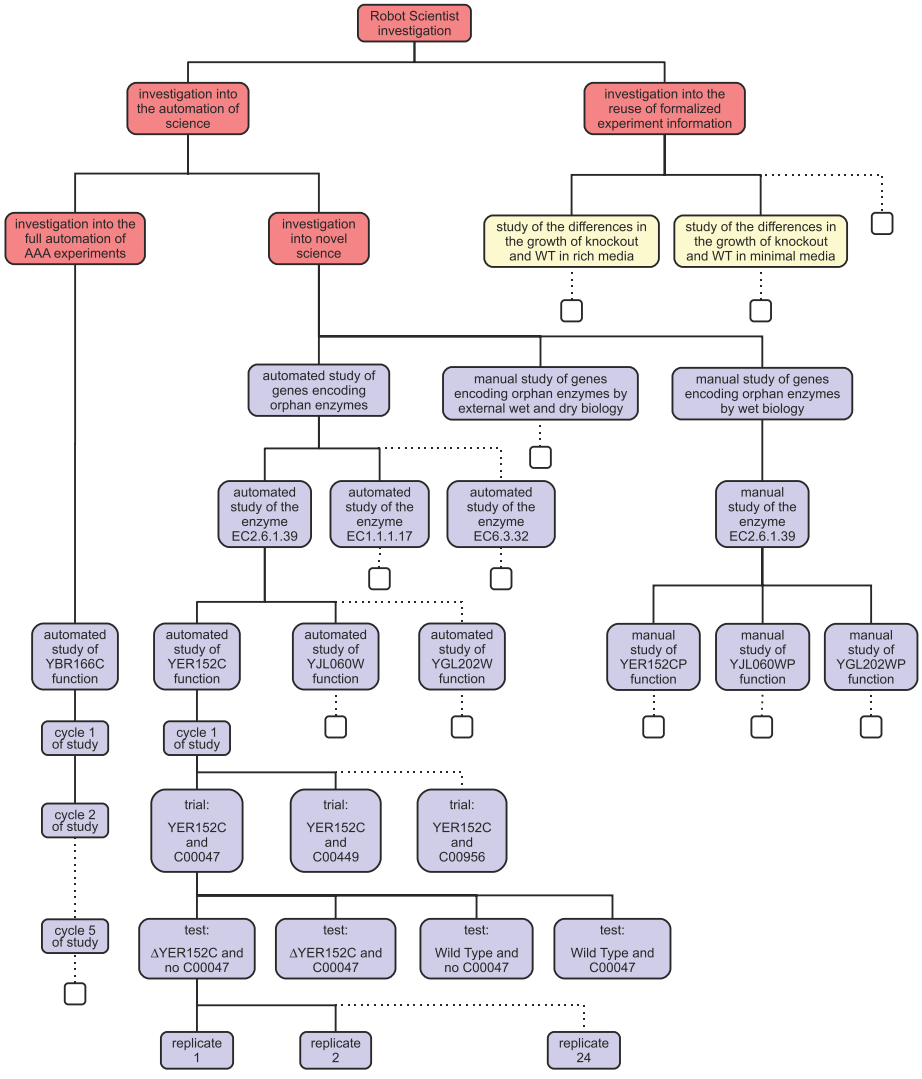

Fig. 3. Structure of the Robot Scientist investigation (a fragment). It consists of two main parts: an investigation into the automation of science and an investigation into the reuse of formalized experiment information. The top levels involve AI research (red), which requires research in functional genomics (blue) and systems biology (yellow). Each level of research unit (studies, cycles, trials, tests, and replicates) is characterized by a specific set of properties (fig. S3) (*16*). Such a nested structure is typical of many scientific experiments, where the testing of a top-level hypothesis requires the planning of many levels of supporting work. What is atypical in Adam's work is the scale and depth of the nesting.

A major motivation for the formalization of experimental knowledge is the expectation that such knowledge is more easily reused to answer other scientific questions. To test this, we investigated whether we could reuse Adam's functional genomic research (*16*). An example question investigated was the relative growth rates (μ_{max}) in rich and defined media of the deletion strains compared with those of the wild type. What was observed, in both media, was a skewed distribution, with a few deletants having a much lower μ_{max} than that of the wild type, but most having a slightly higher μ_{max}. These observations question the common assumption that wild-type *S. cerevisiae* is optimized for μ_{max} and provide quantitative test data for yeast systems biology models (*19*).

It could be argued that the scientific knowledge "discovered" by Adam is implicit in the formulation of the problem and is therefore not novel. This argument that computers cannot originate anything is known as Lady Lovelace's objection (*20*): "The Analytical Engine has no pretensions to *originate* anything. It can do *whatever we know how to order it* to perform" (her italics). We accept that the knowledge automatically generated by Adam is of a modest kind. However, this knowledge is not trivial, and in the case of the genes encoding 2A2OA, it sheds light on, and perhaps solves, a 50-year-old puzzle (*21*).

Adam is a prototype and could be greatly improved. Its hardware and software are "brittle," so although Adam is capable of running for a few days without human intervention, it is advisable to have a technician nearby in case of problems. The integration of Adam's artificial intelligence (AI) software also needs to be enhanced so that it works seamlessly. To extend Adam, we have developed software to enable external users to propose hypotheses and experiments, and we plan to automatically publish the logical descriptions of automated experiments. The idea is to develop a way of enabling teams of human and robot scientists to work together. The greatest research challenge will be to improve the scientific intelligence of the software. We have shown that a simple form of hypothesis-led discovery can be automated. What remain to be determined are the limits of automation.

References and Notes

1. *Nature* **440**, 383 (2006).
2. T. Hey, A. Trefethen, in *Grid Computing: Making the Global Infrastructure a Reality*, **36** 809 (Wiley, New York, 2003).
3. S. Toulmin, in *Encyclopaedia Britannica Deluxe Edition 2004 CD* (Encyclopaedia Britannica UK, London, 2003).
4. T. Berners-Lee, J. Hendler, O. Lassila, *Sci. Am.* **284**, 34 (May 2001).
5. M. Ashburner *et al.*, *Nat. Genet.* **25**, 25 (2000).
6. R. D. King *et al.*, *Nature* **427**, 247 (2004).
7. B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, in *Machine Intelligence*, B. Meltzer, D. Michie, Eds. (Edinburgh Univ. Press, Edinburgh, UK, 1969), vol. 5, pp. 209–254.
8. P. Langley, H. A. Simon, G. L. Bradshaw, J. M. Zytkow, *Scientific Discovery: Computational Explorations of the Creative Process* (MIT Press, Cambridge, MA, 1987).

9. J. M. Zytkow, J. Zhu, A. Hussam, in *Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, 889 (AAAI Press, Menlo Park, CA, 1990).
10. L. Hood, J. R. Heath, M. E. Phelps, B. Lin, *Science* **306**, 640 (2004).
11. L. N. Soldatova, R. D. King, *J. R. Soc. Interface* **3**, 795 (2006).
12. L. N. Soldatova, A. Clare, A. Sparkes, R. D. King, *Bioinformatics* **22**, e464 (2006).
13. R. D. King *et al.*, The Robot Scientist Project, www.aber.ac.uk/compsci/Research/bio/robotsci/ (2008).
14. J. Warringer, E. Ericson, L. Fernandez, O. Nerman, A. Blomberg, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15724 (2003).
15. K. E. Whelan, R. D. King, *BMC Bioinformatics* **9**, 97 (2008).
16. Materials and methods are available as supporting material on *Science* Online.
17. I. Horrocks, P. F. Patel-Schneider, F. van Harmelen, *Web Semantics* **1**, 7 (2003).
18. J. D. Ullman, *Principles of Database and Knowledge-Base Systems, Vol. I: Classical Database Systems* (Computer Science, New York, 1988).
19. M. J. Herrgard *et al.*, *Nat. Biotechnol.* **26**, 1155 (2008).
20. A. Turing, *Mind* **236**, 433 (1950).
21. T. M. Zabriskie, M. D. Jackson, *Nat. Prod. Rep.* **17**, 85 (2000).
22. M. Young, *J. Exp. Bot.* **24**, 1172 (1973).

23. This work was funded by grants from the Biotechnology and Biological Sciences Research Council to R.D.K., S.G.O., and M.Y.; by a SRIF 2 award to R.D.K.; by fellowships from the Royal Commission for the Great Exhibition of 1851, the Engineering and Physical Sciences Research Council, and the Royal Academy of Engineering to A.C.; and by a RC-UK Fellowship to L.N.S. We thank D. Struttman for help with yeast and M. Benway for help with Adam.

Supporting Online Material

www.sciencemag.org/cgi/content/full/324/5923/85/DC1
Materials and Methods
Figs. S1 to S3
Table S1
References

8 September 2008; accepted 10 February 2009
10.1126/science.1165620

Priming in Systemic Plant Immunity

Ho Won Jung,[1] Timothy J. Tschaplinski,[2] Lin Wang,[3]* Jane Glazebrook,[3] Jean T. Greenberg[1]†

Plants possess inducible systemic defense responses when locally infected by pathogens. Bacterial infection results in the increased accumulation of the mobile metabolite azelaic acid, a nine-carbon dicarboxylic acid, in the vascular sap of *Arabidopsis* that confers local and systemic resistance against the pathogen *Pseudomonas syringae*. Azelaic acid primes plants to accumulate salicylic acid (SA), a known defense signal, upon infection. Mutation of the *AZELAIC ACID INDUCED 1 (AZI1)* gene, which is induced by azelaic acid, results in the specific loss of systemic immunity triggered by pathogen or azelaic acid and of the priming of SA induction in plants. Furthermore, the predicted secreted protein AZI1 is also important for generating vascular sap that confers disease resistance. Thus, azelaic acid and AZI1 are components of plant systemic immunity involved in priming defenses.

Whole plant immunity, called systemic acquired resistance (SAR), often develops after localized foliar infections by diverse pathogens. In this process, leaves distal to the localized infection become primed to activate a stronger defense response upon secondary infection (*1*). Leaves infected with SAR-inducing bacteria produce vascular sap, called petiole exudate, which confers disease resistance to previously unexposed (naïve) plants (*2*, *3*). This indicates that a mobile systemic signal(s) is involved in SAR (*4*). Although the hormone jasmonic acid (JA) accumulates to a high level in petiole exudates from leaves infected with SAR-inducing bacteria, JA does not seem to be the critical signal for SAR (*5*, *6*). Instead, SAR and the production of active exudates require the DIR1 protein, a predicted secreted protein and putative signal carrier in the lipid transfer protein family, and other proteins involved in glycerolipid biosynthesis (*2*, *3*, *7*). Additionally, SAR and exudate-induced resistance appears to require the phenolic metabolite salicylic acid (SA) (*3*, *8*) and possibly methylsalicylate (MeSA) and its methyl

[1]Department of Molecular Genetics and Cell Biology, The University of Chicago, 1103 East 57th Street EBC410, Chicago, IL 60637, USA. [2]Oak Ridge National Laboratory, Environmental Sciences Division, Oak Ridge, TN 37831–6341, USA. [3]Department of Plant Biology, Microbial and Plant Genomics Institute, University of Minnesota, 1500 Gortner Avenue, St. Paul, MN 55108, USA.

*Present address: Boyce Thompson Institute for Plant Research, Tower Road, Ithaca, NY 14853–1801, USA.
†To whom correspondence should be addressed. E-mail: jgreenbe@midway.uchicago.edu

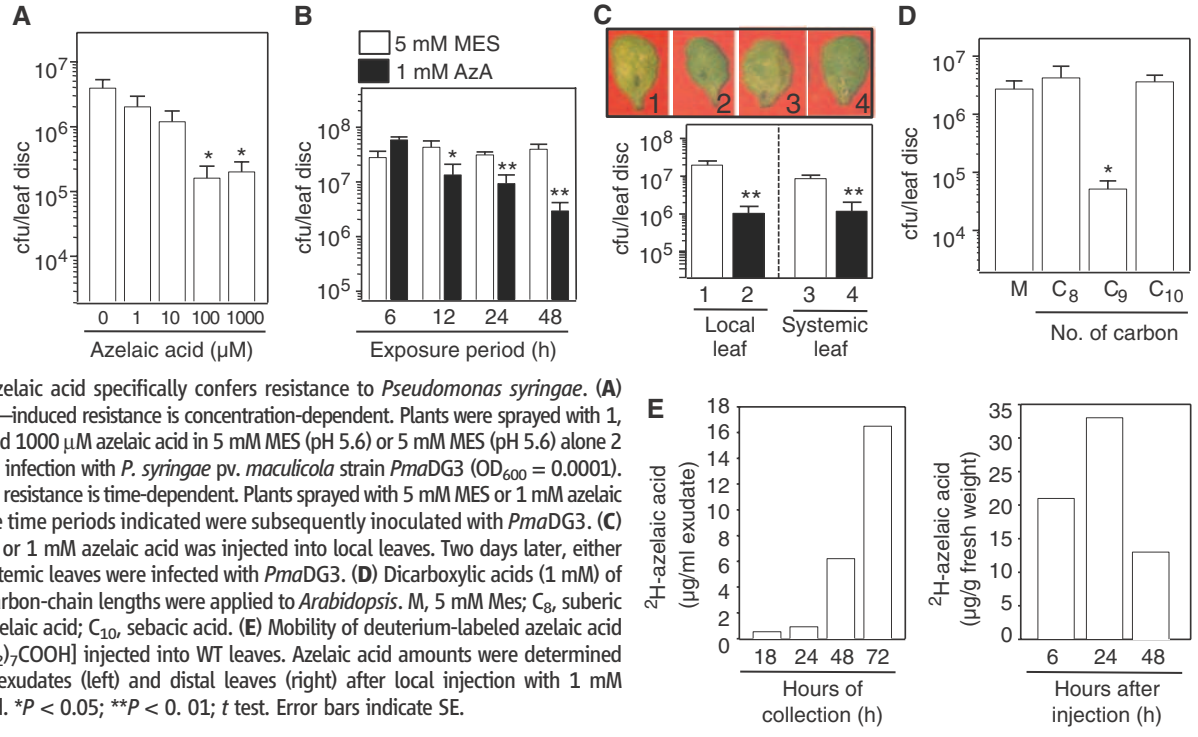

Fig. 1. Azelaic acid specifically confers resistance to *Pseudomonas syringae*. (**A**) Azelaic acid–induced resistance is concentration-dependent. Plants were sprayed with 1, 10, 100, and 1000 μM azelaic acid in 5 mM MES (pH 5.6) or 5 mM MES (pH 5.6) alone 2 days before infection with *P. syringae* pv. *maculicola* strain *Pma*DG3 (OD_{600} = 0.0001). (**B**) Induced resistance is time-dependent. Plants sprayed with 5 mM MES or 1 mM azelaic acid for the time periods indicated were subsequently inoculated with *Pma*DG3. (**C**) 5 mM MES or 1 mM azelaic acid was injected into local leaves. Two days later, either local or systemic leaves were infected with *Pma*DG3. (**D**) Dicarboxylic acids (1 mM) of different carbon-chain lengths were applied to *Arabidopsis*. M, 5 mM Mes; C_8, suberic acid; C_9, azelaic acid; C_{10}, sebacic acid. (**E**) Mobility of deuterium-labeled azelaic acid [$HOOC(CD_2)_7COOH$] injected into WT leaves. Azelaic acid amounts were determined in petiole exudates (left) and distal leaves (right) after local injection with 1 mM azelaic acid. *$P < 0.05$; **$P < 0.01$; *t* test. Error bars indicate SE.