a single verified report whereby breeding or radiation and/or chemical mutagenesis resulted in a toxin, allergen or other hazard that was not known to exist before. These facts support the conclusion that DNA insertions and other types of mutations do not pose unreasonable risks to the environment or to human and animal health, regardless of how they came about.

COMPETING FINANCIAL INTERESTS
The authors declare competing financial interests: details are available at http://www.nature.com/doifinder/10.1038/nbt.2347.

Wayne A Parrott[1], Joseph M Jez[2] & L Curtis Hannah[3]

[1]University of Georgia, Athens, Georgia, USA. [2]Washington University, St. Louis, Missouri, USA. [3]University of Florida, Gainesville, Florida, USA. e-mail: wparrott@uga.edu

1. Waltz, E. *Nat. Biotechnol.* **30**, 215–217 (2012).
2. Ossowski, S. *et al. Science* **327**, 92–94 (2010).
3. Ahloowalia, B.S., Maluszynski, M. & Nichterlein, K. *Euphytica* **135**, 187–204 (2004).
4. Balyan, H.S., Sreenivasulu, N., Riera-Lizarazu, O., Azhaguvel, R. & Kianian, S.F. in *Advances in Agronomy*, Vol. 98 (ed. Sparks, D.L.) 357–414 (2008).
5. Tajima, F. *Genetics* **105**, 437–460 (1983).
6. Huan, N.V., Sugimoto, H. & Harada, K. *Breed. Sci.* **55**, 441–446 (2005).
7. Ross-Ibarra, J., Morrell, P.L. & Gaut, B.S. *Proc. Natl. Acad. Sci. USA* **104**, 8641–8648 (2007).
8. Tenaillon, M.I. *et al. Proc. Natl. Acad. Sci. USA* **98**, 9161–9166 (2001).
9. Pariza, M.W. & Cook, M. *Regul. Toxicol. Pharmacol.* **56**, 332–342 (2010).
10. Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J. & van der Knaap, E. *Science* **319**, 1527–1530 (2008).
11. Rodriguez, G.R. *et al. Plant Physiol.* **156**, 275–285 (2011).

Broad consent in biobanking

To the Editor:
The Feature in the February issue by Scott *et al.*[1] on the policy challenges of biobanking characterizes broad specimen donor informed consent as "ethically contentious." A survey of public attitudes is cited. This same survey found that a significant percentage of individuals are prepared "to consent broadly to future research use and to forego additional choices as a result"[2].

With our perspectives in patient advocacy or at research centers aimed at bringing new regenerative therapies to patients, we have consistently emphasized the value of research donors' perspectives. In the context of protocols for creating immortalized cell lines for banking and distribution, we have also witnessed support for broad consent. Indeed, enthusiasm is even more pronounced among those touched by disease, and patient donors actually express concern that study-specific consent can be burdensome and impede research.

This experience suggests to us that broad consent is ethically responsible, provided there is comprehensive oversight and a robust informed consent process. With the continued support of donors, we look forward to applying this model in biobanking efforts.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Chris Hempel[1], Geoffrey Lomax[2] & Steve Peckman[3]

[1]Addi & Cassi Fund, Reno, Nevada, USA. [2]California Institute for Regenerative Medicine, San Francisco, California, USA. [3]Eli and Edythe Broad Center of Regenerative Medicine, University of California, Los Angeles, California, USA. e-mail: glomax@cirm.ca.gov

1. Scott, C.T. *et al. Nat. Biotechnol.* **30**, 141–147 (2012).
2. Simon, C.M. *et al. Genet. Med.* **13**, 821–831 (2011).

An accelerated workflow for untargeted metabolomics using the METLIN database

To the Editor:
Metabolites, which are typically recognized as small molecules that are involved in cellular reactions, provide a functional signature of phenotype that is complementary to the upstream biochemical information obtained from genes, transcripts and proteins. The high correlation between metabolites and phenotype has created a surge of interest in the field that is reflected in the number of metabolomic publications growing from just a few articles in 1999 to over 5,000 in 2011. Although relatively new compared with its genomic and proteomic predecessors, research in the field of metabolomics has already led to the discovery of biomarkers for disease, fundamental insights into cellular biochemistry and clues related to disease pathogenesis[1,2].

The success of metabolomics over the past decade has relied largely on advances in mass spectrometry instrumentation, which make it possible to detect thousands of metabolites simultaneously from a biological sample. Coupled with developments in bioinformatic tools such as XCMS Online (https://xcmsonline.scripps.edu/)[3], it has now become relatively routine to comprehensively compare the intensities of thousands of metabolite peaks in one sample group to those in another in an untargeted manner. This approach, called untargeted metabolomics, has the potential to implicate unexpected pathways with a unique phenotype or disease process.

Despite the attractiveness of having a comprehensive and unbiased approach for profiling metabolites that is analogous to those used in the other 'omic' sciences, an overwhelming proportion of the metabolomic community exclusively uses a targeted platform in which only a specified list of metabolites is measured. The benefit of such a targeted platform is speed. Unlike the untargeted platform, after the targeted mass spectrometry methods are established, minimal effort and resources are required to profile these specific metabolites over a large number of samples. In contrast, the major bottleneck of untargeted metabolomics has been the challenge of determining the identities of the peaks found to be dysregulated in the untargeted profiling data.

Traditionally, the untargeted metabolomic platform involves multiple steps (**Fig. 1**). The first step is acquiring global mass spectrometry data for each of the samples. Next, these data are analyzed using bioinformatic software that performs quantitative analyses to find peaks that are significantly different between sample groups. The investigator then typically searches the mass-to-charge (m/z) ratios of the peaks of interest manually in metabolite databases. Searches that return hits within the mass accuracy of the instrument are considered to be putative identifications. To confirm the identifications, tandem mass spectrometry (MS/MS) data from the research sample are then compared to the MS/MS data of a commercial standard. To obtain the MS/MS data, a targeted MS/MS analysis is typically performed on one of

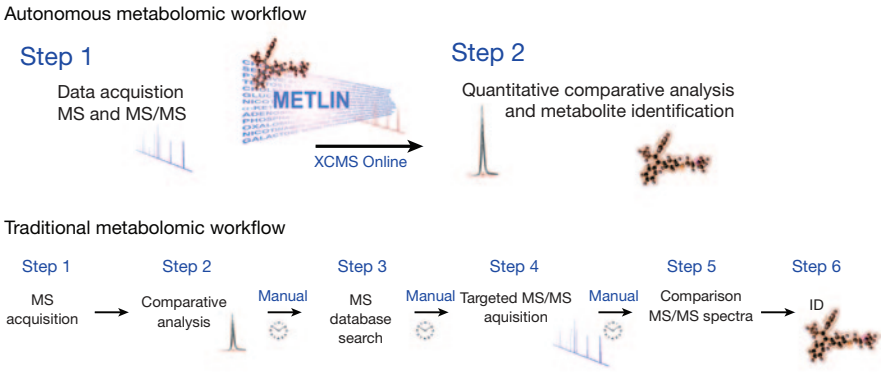

Figure 1 Schematic representation of the traditional metabolomic workflow involving six steps and the new METLIN-based workflow with only two steps. In the two-step autonomous workflow, mass spectrometry (MS) and MS/MS data are acquired simultaneously during profiling and searched in the METLIN database for automated identification, thereby reducing the time of the workflow from days or weeks to minutes or hours.

the research samples for which the peak was determined to be upregulated. The fragmentation pattern of the MS/MS data is then manually compared with that of the MS/MS data from a commercial standard (however, not all commercial standards, for example, stereoisomers, can be resolved by MS/MS data alone).

To facilitate identification of metabolites in the untargeted workflow, we launched a freely accessible metabolite database called METLIN[4] in 2004 (http://metlin.scripps.edu/) that incorporates MS/MS data from model compounds. Recently, other metabolite databases such as the Human Metabolome Database (HMDB)[5], MassBank[6] and LipidMaps[7] have also begun incorporating MS/MS data for standard compounds. These repositories allow investigators to compare MS/MS data from their research samples to MS/MS data from model compounds catalogued in the database and thereby improve the speed, efficiency and cost effectiveness of untargeted studies.

Over the past 7 years, our objective has been to generate a sufficiently large MS/MS data library that can be used in an automated manner to revise the traditional untargeted metabolomic workflow (**Fig. 1**). Since we originally reported the establishment of METLIN in 2005, we have increased the number of MS/MS spectra that are included in the database by a factor of 150. As of April 2012, METLIN contains MS/MS data on >10,000 distinct metabolites at four different collision energies. These data were collected using an electrospray ionization/quadrupole time-of-flight (ESI-QTOF) mass spectrometer in both the positive and negative detection modes, representing a total number of >48,000 high-resolution spectra. To estimate the current coverage of physiologically relevant metabolites in METLIN and the other three largest databases available (HMDB, MassBank and LipidMaps), metabolites were isolated from *Escherichia coli* and standard human serum using defined protocols[8]. Samples were analyzed in both the positive and negative modes with an ESI-QTOF mass spectrometer (**Supplementary Methods**). Each peak detected (excluding isotopes) was searched in each of the four databases. **Figure 2** shows the number of hits for each database and also the subset of the hits for which MS/MS data were available to confirm the metabolite identification.

In addition to its increased size, here we describe a new version of the METLIN database that has advanced functionality to automate metabolite identification and reduce the labor-intensive bottleneck that has traditionally been associated with untargeted metabolite profiling. Instead of manually comparing the MS/MS data from research samples to the MS/MS data of commercial standards, the new version of METLIN allows metabolomic investigators to upload their MS/MS data to the METLIN database so the comparisons can be performed in an automated way. By using automated MS/MS data matching, metabolite identities can be confirmed much more efficiently and quickly compared with the traditional untargeted metabolomic workflow. The quality of the match between the MS/MS data from the research sample and the MS/MS data from the METLIN library is measured by a newly introduced METLIN scoring system, which is based on a modified version of the established X-Rank scoring system[9]. To evaluate the correlation of METLIN MS/MS data to MS/MS data acquired using different instrument platforms, we performed a comparative experiment was performed using 23 metabolite standards. The compounds were measured on five different instruments, and the resulting spectra were matched against the METLIN database. Based on the modified X-Rank scoring system, the correct result was returned as the first hit for 90 out of the 101 spectra (89.1%; **Supplementary Table 1** and **Supplementary Figs 1–27**).

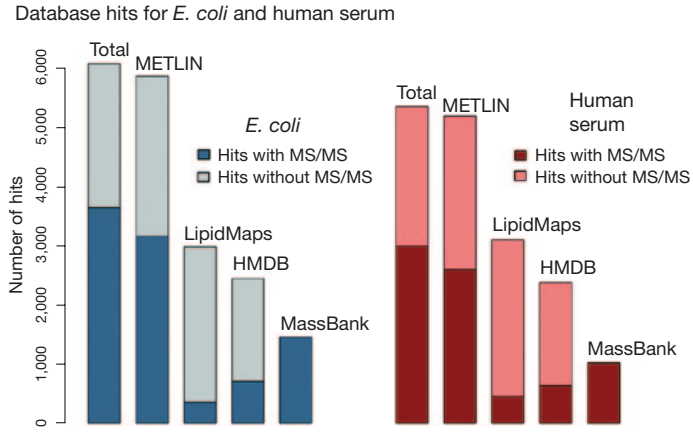

Figure 2 Estimate of the physiological relevance of metabolite coverage in metabolomic databases. Metabolites from human serum and *E. coli* were isolated and analyzed in both the positive and negative modes by ESI-QTOF mass spectrometry, and the mass of each metabolite was searched with a tolerance of 5 parts per million in the METLIN, LipidMaps, HMDB and MassBank databases. LipidMaps contains data primarily on lipids, which is only a subset of the metabolome, but was included in the comparison for the sake of completeness. For human serum, 12,170 features were detected and searched, and for *E. coli*, 11,641 features were detected and searched. The number of hits on the basis of accurate mass are shown in light blue and light red for *E. coli* and human serum, respectively. The subset of those hits that also contained MS/MS data are shown in dark blue and dark red for *E. coli* and human serum, respectively.

Some classes of metabolites produce characteristic fragments or neutral losses in their MS/MS spectra that can be used as signatures for unique chemical functional groups. For example, the MS/MS spectra of phosphatidylcholines are characterized by a fragment at *m*/*z* 184.07. For instances in which the MS/MS data uploaded by a user do not match any compound in the database, the new version of the METLIN database will search the MS/MS data for characteristic fragments that can be used for molecular classification. The search can also be performed manually by accessing the 'fragment search' or 'neutral loss search' options. These tools provide a new mechanism by which unknown metabolites can be chemically classified, and they take advantage of the large amount of MS/MS data in the library.

To highlight the new database functionalities, we performed MS/MS on select peaks from the metabolite extracts of *E. coli* and human serum. These data were uploaded to the METLIN database, and fragment matching was performed using the automated feature described above. Representative examples of metabolites identified on the basis of the mass spectrometry and MS/MS data using this method are shown in **Supplementary Figures 28–32**. The compounds identified ranged from lipids to smaller, polar metabolites. Additionally, representative examples of unknown compounds that were classified by characteristic fragments are also shown.

With the combination of the METLIN functionalities described here and the increasing speed of QTOF instrumentation for performing MS/MS, there is the potential to reduce the untargeted metabolomic workflow to just two steps (**Fig. 1**). Using high–scan-speed QTOF instruments, mass spectrometry and MS/MS data can be acquired simultaneously in a single run. Quantitative information can then be extracted from the data using the bioinformatic software XCMS Online, and metabolites can be identified simultaneously by matching the MS/MS data with MS/MS data in the METLIN database in an automated fashion, an approach that is self-directed or autonomous in nature.

With this truncated workflow, the time needed to perform untargeted profiling and the subsequent metabolite identification may be reduced to minutes or hours as compared to the days or weeks needed with the traditional workflow. The results shown here from automated MS/MS matching highlight the applicability of the method for performing high-throughput, untargeted metabolomics using this type of accelerated workflow. Moreover, we have shown that the coverage of the METLIN database enables the characterization and identification of thousands of naturally occurring metabolites in biological samples. Thus, the new METLIN database has the potential to expedite the workflow for untargeted metabolomics as more investigators obtain mass spectrometry instrumentation that can produce high-quality MS/MS data with increasing speed and sensitivity.

Note: Supplementary information is available at http://www.nature.com/doifinder/10.1038/nbt.2348.

ACKNOWLEDGMENTS

This work was supported by the California Institute of Regenerative Medicine (TR1-01219), the US National Institutes of Health (R24 EY017540, P30 MH062261, RC1 HL101034 and P01 DA026146) and the US National Institutes of Health National Institute on Aging L30 AG0 038036 (G.J.P.). Financial support was also received from the US Department of Energy (grants FG02-07ER64325 and DE-AC0205CH11231).

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Ralf Tautenhahn[1,2], Kevin Cho[1,2], Winnie Uritboonthai[1,2], Zhengjiang Zhu[1,2], Gary J Patti[3–5] & Gary Siuzdak[1,2]

[1]Department of Chemistry, Center for Metabolomics, The Scripps Research Institute, La Jolla, California, USA. [2]Department of Molecular Biology, The Scripps Research Institute, La Jolla, California, USA. [3]Department of Chemistry, Washington University School of Medicine, St. Louis, Missouri, USA. [4]Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA. [5]Department of Medicine, Washington University School of Medicine, St. Louis, Missouri, USA.
e-mail: gjpattij@wustl.edu or siuzdak@scripps.edu

1. Yanes, O. *et al. Nat. Chem. Biol.* **6**, 411–417 (2010).
2. Patti, G.J. *et al. Nat. Chem. Biol.* **8**, 232-234 (2012).
3. Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R. & Siuzdak, G. *Anal. Chem.* **78**, 779–787 (2006).
4. Smith, C.A. *et al. Ther. Drug Monit.* **27**, 747–751 (2005).
5. Wishart, D.S. *et al. Nucleic Acids Res.* **37**, D603–D610 (2009).
6. Horai, H. *et al. J. Mass Spectrom.* **45**, 703–714 (2010).
7. Sud, M. *et al. Nucleic Acids Res.* **35**, D527–D532 (2007).
8. Yanes, O., Tautenhahn, R., Patti, G.J. & Siuzdak, G. *Anal. Chem.* **83**, 2152–2161 (2011).
9. Mylonas, R. *et al. Anal. Chem.* **81**, 7604–7610 (2009).

Successful suppression of a field mosquito population by sustained release of engineered male mosquitoes

To the Editor:

Our paper published last year described the results of preliminary release experiments showing that engineered sterile male mosquitoes could mate with females in a wild population in the Cayman Islands[1]. This trial was supported by simple simulation models indicating that sustained release of sufficient numbers of such males should substantially suppress a target population within a few weeks or months[2–4]. In the following letter, we describe a field release experiment testing this proposition.

The sterile insect technique is an environmentally friendly, species-specific method of pest control that is used to successfully control several agricultural pest insects[5]. Large numbers of sterile insects are released to mate with their wild counterparts and thereby reduce their reproductive potential. However, despite its attractive features, this technique is not in operational use against mosquitoes, in part because of damaging effects of sterilizing doses of radiation on the released mosquitoes[6–8]. Following a similar principle, we have proposed that engineered males carrying a dominant lethal transgene could be released to mate with wild females; the resulting progeny would die as a result of the lethal effect of the transgene. We named this system RIDL (release of insects carrying a