

The Addgene repository: an international nonprofit plasmid and data resource

Joanne Kamens*

Addgene, Cambridge, MA 02139, USA

Received August 01, 2014; Revised September 12, 2014; Accepted September 16, 2014

ABSTRACT

The Addgene Repository (<http://www.addgene.org>) was founded to accelerate research and discovery by improving access to useful, high-quality research materials and information. The repository archives plasmids generated by scientists, conducts quality control, annotates the associated data and makes the plasmids and their data available to the scientific community. Plasmid associated data undergoes ongoing curation by members of the scientific community and by Addgene scientists. The growing database contains information on >31 000 unique plasmids spanning most experimental biological systems and organisms. The library includes a large number of plasmid tools for use in a wide variety of research areas, such as empty backbones, lentiviral resources, fluorescent protein vectors and genome engineering tools. The Addgene Repository database is always evolving with new plasmid deposits so it contains currently pertinent resources while ensuring the information on earlier deposits is still available. Custom search and browse features are available to access information on the diverse collection. Extensive educational materials and information are provided by the database curators to support the scientists that are accessing the repository's materials and data.

INTRODUCTION

Plasmids are circular fragments of double-stranded DNA that can replicate independently of chromosomal DNA. In 1952, Joshua Lederberg was the first scientist to give this name to small, extra-chromosomal, heritable determinants (1). By the 1970s and the advent of molecular biology tools, such as restriction enzymes and DNA ligases, scientists were able to use plasmids to study individual genes. Today scientists study genes and other genetic elements using a multitude of specifically engineered plasmids. Plasmids have become a ubiquitous research tool fulfilling a large num-

ber of functions for many reasons. First, steady advances in cloning methods have made it increasingly easy to generate useful and diverse plasmid constructs. Second, plasmid DNA is stable and, once introduced to bacteria, can be stored and indefinitely replicated. Finally, plasmids are useful for studies on virtually any biological pathway or system in a wide variety of organisms. For all these same reasons, plasmids are excellent research materials for archiving and sharing.

As grant funding becomes harder to access, scientists are finding it increasingly important to be efficient with resources (2,3). Time and money are wasted in regenerating similar plasmid constructs in different laboratories. It makes sense for useful plasmids to be shared widely to accelerate research. In the past, plasmids were generated by graduate students and postdoctoral candidates then used for a few or even one publication before being relegated to the freezer when the trainee moved on to their next position. Valuable materials were languishing unused and potential collaborations were never initiated. The Addgene Repository and other plasmid repositories prevent this waste by helping scientists to archive and share plasmids (4–6). The Addgene Repository and database are focused on encouraging materials sharing and collaboration in addition to careful archiving of samples (7). Search functions are improved to make sure that the information on all materials can be discovered by researchers. New deposits are always being added to the database to ensure that the materials remain relevant for current research. That being said, some of the most popular plasmids have been in the database for over 10 years (8).

Repositories also play a role in standardization of reagents as an important part of improving data reproducibility (9–11). Most journals and funding agencies suggest or require that scientists make research materials available to other scientists after publication. It is not easy for every lab to comply with this best practice. Materials sharing, especially for international collaboration, can be hampered by logistical constraints. Repositories provide tested, standardized reagents and associated data to help scientists confirm and extend published results.

*To whom correspondence should be addressed. Tel: +1 617 225 9000 (Ext 112); Fax: +1 888 734 0533; Email: joanne.kamens@addgene.org

OVERVIEW OF THE PLASMID DATABASE

Growth and curation of the database

More than 31,000 unique plasmids are available for request from the Addgene database. Deposits come to the repository by a few mechanisms. The Addgene Scientist Team scans the literature and solicits the deposit of published plasmids. Addgene Outreach Scientists visit laboratories around the world to meet as many scientists as possible in person. Finally, as more laboratories join the Addgene Repository community and are familiar with the advantages of using a data and materials repository, an increasing number of contributions are initiated by depositing scientists themselves.

The Addgene Repository archives plasmids and plasmid collections that may be published, unpublished or published only in part (some example plasmid collections are listed in Table 1; see for example, the Davidson FP collection). In the process of generating data for publications or grant applications, current cloning methods often result in useful intermediate constructs or plasmids that would not be used in a publication but are useful to others and merit database inclusion.

Relevant plasmid information, such as cloning sites, mutations, resistance markers and available sequence information, is stored in the database. If applicable, plasmids are linked to the article in which they were first described, which provides relevant information for requesting scientists and ensures that depositing labs are cited appropriately. Supporting data and documents can also be provided for each plasmid. For example, a plasmid encoding a fluorescent protein might have associated data files with example images of cellular localization. These associated data are a valuable component of the database for scientists planning experiments.

After data entry, every plasmid undergoes sequence analysis of key regions (ends, tags, mutations, etc.) for construct validation and quality control. Any discrepancies from the depositor-provided sequences are resolved before the plasmid cloning data and sequences are posted to the database. Samples are stored in triplicate (including storage of the full library at an offsite back-up facility). The technology transfer process largely goes on behind the scenes so that depositing and requesting scientists can receive the benefits of having the correct agreements in place without having to spend a significant amount of time for each exchange.

Scientists are able to initiate a deposit before a planned paper is submitted. All of the information provided for the deposit can be embargoed, that is, not made publicly visible until publication. Pre-publication deposit is a valuable benefit to scientists. As soon as the deposit is initiated each plasmid is assigned a unique Addgene Accession number which can be included in the materials and methods section of a publication. Similar to a reagent catalog number, inclusion of an Addgene Accession number in the publication helps readers to quickly find plasmid cloning details. On the Pubmed database of indexed citations and abstracts (<http://www.ncbi.nlm.nih.gov/pubmed>), the abstract listings for publications containing Addgene Repository plasmid references include a direct link to the webpage displaying the

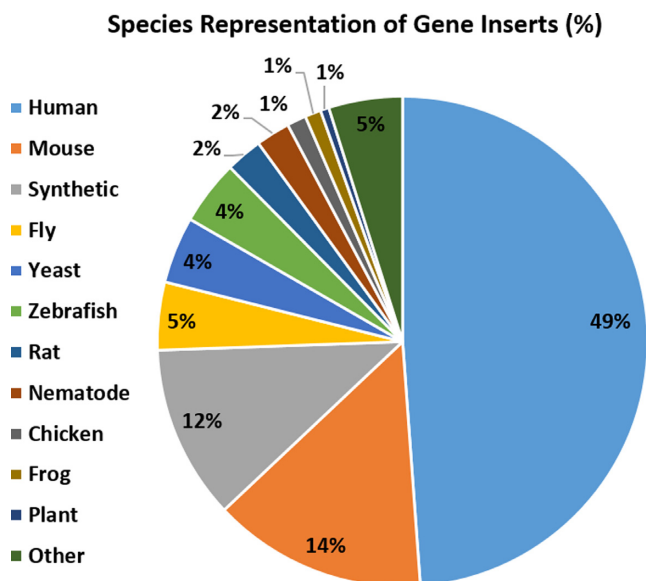


Figure 1. Species Representation of Gene Inserts. As of July 2014 the proportion of plasmids in the Addgene library with inserts from each of the indicated species. Insert species data were provided by depositors or available for 22,825 plasmids in the library. For some plasmids (e.g. research tools, synthetic genes) there is no appropriate inserts species.

plasmids deposited for that paper. Easy access to this page and detailed plasmid information adds value to online publications.

The plasmid library

Plasmids in the database are useful for studies in a wide variety of different experimental systems and model organisms. This diversity reflects the research fields of the >2000 depositing laboratories from >500 institutions around the world. Almost half of the library's plasmids have been designated as containing human gene sequence inserts (Figure 1). Most fields of research are represented in the library. Some areas for which the repository has especially rich resources include: neuroscience, stem cell research, lentiviral transduction, fluorescent protein tagging, optogenetics and genome engineering. A review of the 25 most accessed plasmids over time reflects popular plasmid technologies that are all well represented in the library (12). This top 25 list includes plasmids for retroviral expression, lentiviral expression, luciferase reporter assays, RNA interference studies, generation of induced pluripotent stem cells, mammalian Cre-Lox control and TALEN and CRISPR plasmids for genome engineering.

The library is curated into convenient collections to allow requesting scientists to more easily browse plasmids for experimental design and planning (Table 1). For example, one of the most popular collections pages is the Plasmid Backbone Guide (addgene.org/empty-backbones), which helps requestors choose among the library's 2,500 empty backbones (Figure 2). Examples of other highly accessed pages include the Fluorescent Protein Guide (addgene.org/fluorescent-proteins), Lentiviral Plasmid Collection (addgene.org/lentiviral), Popular Plasmids

Table 1. A selection of plasmid resource pages available on the Addgene Repository website

Resource	URL	Description
BioBricks™ Foundation	addgene.org/biobrick_public_agreement_collection	Plasmids contributed under the BioBrick™ Public Agreement
Biosensors	addgene.org/fluorescent_proteins/biosensors	Small molecule and gene-specific biosensors
Cell Migration Consortium	addgene.org/cmc	Useful plasmids for cell migration research
CRISPR plasmids	addgene.org/CRISPR	CRISPR genome editing plasmids by function
CRISPR reference	addgene.org/CRISPR/reference	Supporting technologies, protocols, videos and other educational support
Davidson FP Collection	addgene.org/fluorescent_proteins/davidson	Empty backbones, FP-tagged ORFs including excitation/emission, localization, sequence data
Fluorescent Protein Guide	addgene.org/fluorescent_proteins	Guide to Addgene Fluorescent Protein collections
FRET Guide	addgene.org/fluorescent_proteins/fret	Empty vectors and reference standards for Förster resonance energy transfer
Genome Engineering Guide	addgene.org/genome_engineering	Introduction to genome engineering resources at Addgene
KLF Plasmid Resource	addgene.org/KLF	Krüppel-like factors: family of DNA binding transcriptional regulators
Lentiviral Plasmid Collection	addgene.org/lentiviral	Lentivector lists including popular plasmids in this category
Lentiviral Protocols & Resources	addgene.org/lentiviral/protocols-resources	Instructional resources including a webinar, protocols and definitions
Mammalian RNAi Tools	addgene.org/mammalianrna	Empty vectors to clone shRNAs for gene silencing studies
Open Source Wnt Project	addgene.org/wnt/Open_Source_Wnt	Plasmids to study secreted morphogens that control developmental processes
Optogenetics	addgene.org/optogenetics	Optogenetic tools: actuators and sensors
Parkinson's Disease Plasmid Resource	addgene.org/PD	Assembled by the Michael J. Fox Foundation for Parkinson's Disease and individual laboratories
Plasmid Backbones	addgene.org/empty_backbones	Guide to choosing an appropriate vector backbone for your construct
Popular Plasmids	addgene.org/popular-plasmids	Guide to some of the most popular tools in the library
Structural Genomics Consortium	addgene.org/sgc	SGC plasmids for determining 3-D protein structures
Subcellular Localization	addgene.org/fluorescent_proteins/localization	Guide to backbones with subcellular localization elements
TALEN Plasmids	addgene.org/TALEN	Kits available for TALEN assembly
TALEN Reference	addgene.org/TALEN/guide	Overview of TALEN technologies
Zinc Finger Consortium	addgene.org/zfc	The ZFC promotes continued research on engineered zinc finger technology

Empty Backbone Plasmids at Addgene (#)

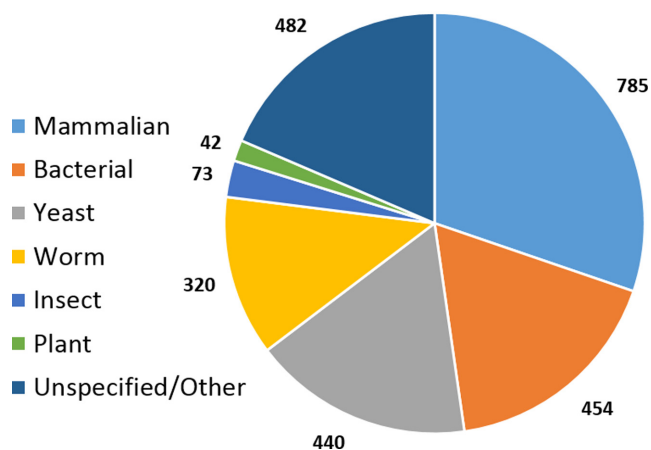


Figure 2. Empty Backbone Plasmids at Addgene. The number of backbone plasmid vectors in the Addgene library designated by depositor recommended expression. As of July 2014 the total number of plasmids designated as 'Empty Backbone' is 2,596.

(addgene.org/popular-plasmids) and Genome Engineering Guide (addgene.org/genome_engineering).

Research consortia also work with the repository to generate customized consortium home pages to make their plasmid collections as accessible as possible to scientists in their fields. See, for example, the Biobrick Public Agreement Collection (addgene.org/biobrick_public_agreement_collection), the Parkinson's Disease Plasmid Resource (addgene.org/PD) and the Structural Genomics Consortium Collection (addgene.org/sgc). The consortium pages offer collaborative research programs a way to highlight their resources and to encourage deposit to the database.

ACCESSING THE ADDGENE DATABASE

Search features

Addgene Database queries start with the search box in the top right corner of every page at addgene.org. Scientists can search by gene name, gene alias, depositing scientist, publication author, Addgene Accession Number or any other relevant search term. The search results page displays two separate search boxes. The 'Full Site Search' box searches all of the site's web content. This search can be used, for example,

to find information about ordering, educational resources or to search for a publication author that is not the Principal Investigator (PI). The 'Search for Plasmids' box is optimized to search and retrieve plasmid-specific results. The plasmid search can be refined by toggling 'Show/hide additional search fields' and entering criteria in multiple fields.

If a gene name/alias is used in the 'Search for Plasmids' box, the first results will be associated resource pages if there are matches to the search term. These will be followed by the summary list of matching plasmid records. The summary list entries link to the Plasmid Data Page (see description below). Plasmid search results can be sorted or filtered with check-box toggles to the right of the results. A flame icon indicates if a plasmid has been requested more than 20 (yellow), 50 (red) or 100 (blue) times. This information is provided because the popularity of a plasmid is one indication of its usefulness and level of experimental validation.

If the name of a PI is used in the 'Search for Plasmids' box, the first results will be links to the Lab Plasmid Pages of depositors with that last name. A customized, public Lab Plasmid Page is generated for each depositing PI or lab head. A link back to the depositing lab's webpage on their university site is displayed here. Deposited plasmids can be browsed on the Lab Plasmid Page in lists organized by plasmid name or, alternatively, by the associated publication. HTML code is also provided to make it easy to post a link from the institutional lab website to their corresponding Addgene Lab Plasmid page. This makes it simple for other scientists to find plasmid-associated data.

Plasmid data page

Each plasmid in the collection has a unique Addgene Accession number and Plasmid Data Page (Figure 3). The Plasmid Data Page first displays the depositing laboratory and the associated publication, if applicable. The 'Purpose' field information is provided by depositors or by Addgene Repository curators. It describes, in simple terms, a suggested experimental use for the plasmid. This field can be particularly helpful when scientists are comparing similar reagents to decide on what will work best for their planned experiments. The Plasmid Data Page displays details of the backbone, growth conditions, gene inserts and cloning information for each inserted sequence. Any additional documents that have been provided by the depositing scientist, such as protocols or maps, can be found under 'Resource Information'. Maps for the Plasmid Data Page are constructed based on available sequence information. Some sequences might be partial or insert-only sequence. Actual or predicted full sequence is available for some plasmids. From the 'View all sequences' link on the Plasmid Data page, users can access a simple analysis tool for basic sequence manipulations including feature mapping, sequence download, NCBI 'BLAST' search, sequence alignment, restriction pattern prediction and translation.

At the bottom of the Plasmid Data Page, recommended citation text is displayed. This text can be used to cite the plasmid in the methods and references sections of a publication resulting from experiments done with a requested reagent. Every effort has been made to make plasmid citation as easy as possible to ensure the depositing scientists

are acknowledged when their materials are used in published studies. Citation is required by the terms of every request to acknowledge the depositing scientist. Future requestors find it helpful when the Addgene Accession number is included in the publication. One of the many advantages of sharing materials through a repository is that this sharing has been demonstrated to increase the number of citations of the original publication in which the reagent appeared (13).

PI and Technology Transfer Office (TTO) account pages

Each depositing scientist is provided with access to a password protected PI Account Page. The PI Account Page lists all requests made for that lab's deposited plasmids including the material requested, the date of the request, the requesting PI and their institution. The PI also has the option to receive daily, weekly or monthly email updates with this information. Easy access to plasmid distribution information allows depositors to connect with requesting scientists who may become future collaborators. In addition, depositors can use these data in grant applications to comply with 'resource sharing plan' requirements.



The TTO for each organization is provided with access to a password protected TTO Account Page. This page displays an inventory of all materials deposited by labs at this organization and allows TTOs to manage organizational contacts and reporting options. TTOs can view a log of all plasmids distributed on their behalf. The TTO Account Page also provides information on all plasmid requests fulfilled for their organization through the Addgene Repository.


PLASMID INFORMATION RESOURCES

The Addgene website contains extensive support information for using and requesting materials from the database (for example, addgene.org/faq). Addgene also has a team of molecular biologists available to help scientists find appropriate reagents. As the size of the Addgene database has increased over the years, so have the number of technical service questions. Realizing that the breadth of technical support questions provides good insight into the information scientists need to plan and carry out plasmid cloning projects, the Addgene Repository scientists have responded by creating a substantial amount of educational content for the community.


In addition to the detailed guides for specific collections (Table 1), Addgene Repository scientists have developed broad educational resources and reference information for the molecular biologist (addgene.org/plasmid_reference). The Molecular Cloning Guide includes an ever growing series of protocols for molecular biology procedures including cloning, plasmid analysis, bacterial culture, DNA purification and nucleic acid quantification. Additional protocols for plasmid handling and analysis can be found under the Protocols for Addgene Plasmids section, including how to treat requested plasmids when you receive them. The Molecular Biology Tools and Resources section has basic information, such as a listing of common bacterial strains and a table of common epitope tags. One

[Browse](#) / [George Church](#) / [Mali et al](#) / [hCas9](#)

 **hCas9**
(Plasmid #41815)
 Print



[Enlarge](#) | [View all sequences](#)



PURPOSE
Expresses human codon optimized Cas9 nuclease for genome engineering

DEPOSITING LAB
[George Church](#)

PUBLICATION
[Mali et al Science](#). 2013 Jan 3.

PRICE
\$65 (USD)

Add to Cart

Available to Academic and Nonprofits Only

BACKBONE

Vector backbone: [pcDNA3.3-TOPO](#)
([Search Vector Database](#))

Backbone manufacturer: [Invitrogen](#)

Vector type: [Mammalian Expression, CRISPR](#)

GENE/INSERT

Gene/Insert name: [Cas9](#)

Insert Size (bp): [4140](#)

Mutation: [human codon-optimized](#)

Promoter: [CMV](#)

GROWTH IN BACTERIA

Bacterial Resistance(s): [Ampicillin](#)

Growth Temperature: [37°C](#)

Growth Strain(s): [Stbl3](#)

Copy number: [High Copy](#)

CLONING INFORMATION

Cloning method: [TOPO Cloning](#)

5' sequencing primer: [CMV_F](#)
([Common Sequencing Primers](#))

SEQUENCE INFORMATION

Depositor Sequences: [Full \(1\)](#) [Partial \(2\)](#)

Addgene Sequences: [Partial \(2\)](#)

RESOURCE INFORMATION

Supplemental documents:

- [hCRISPR gRNA Synthesis Protocol](#)

Terms and Licenses:

- [UBMTA](#)

Related items:

From this article

[George Church Lab Plasmids](#)

[CRISPR/Cas plasmids](#)

Commonly requested with:

[gRNA_Cloning Vector](#)

[hCas9_D10A](#)

[gRNA_AAVS1-T2](#)

[gRNA_GFP-T1](#)

[pX330-U6-Chimeric_BB-CBh-hSpCas9](#)

Figure 3. The Addgene Plasmid Page. Each plasmid has a unique information page. See text for details on the different informational sections.

of the most accessed reference guides is a digital collection of over 4000 vector backbones assembled from publications and commercially available sources called Vector Database (addgene.org/vector-database). Vector Database, along with the Addgene Guide to Choosing a Perfect Backbone (addgene.org/empty_backbones), helps scientists find and select plasmid starting points for their plasmid constructs.

In late 2013, the Addgene Repository launched a blog to enable rapid publishing of timely informational and educational articles for the scientific community (blog.addgene.org).

Addgene Scientists write content for the blog, 'A Better Way to Share Science', and excellent contributions are provided by guest bloggers. Topics include highlights of new plasmids in the repository, technical tips on plasmid design, editorial viewpoints on current topics in science, as well as career content of interest to science trainees and their mentors.

OUTCOMES

The Addgene Repository's distribution service can influence uptake and dissemination of new technologies. One example of this is how rapidly labs have been able to design and carry out experiments related to the growing field of CRISPR/Cas9 genome engineering (14,15). Over 25,000 CRISPR/Cas9-associated plasmids have been distributed since the first papers describing the use of this technology were published in late 2012 and early 2013 (16–19). The dedication of this research community to sharing materials and information has accelerated research on this promising technology. A recent search on Google Scholar (<http://scholar.google.com>) for the terms 'addgene AND crispr' returned almost 500 publication results in just 2 years.

The repository distributed >94,000 plasmids in 2013 and is on track to distribute over 115,000 in 2014. The proportion of plasmids distributed internationally has steadily increased such that over half of all requests are now from outside the United States. The Addgene Repository is proud to be helping scientists share plasmids and data in all 50 of the United States and in 79 countries internationally. The mission to improve access to useful research materials includes a focus on helping all scientists share plasmids and data wherever they are located. The richness and scale of the library and database will make it possible for the Addgene Repository to continue to grow its community of collaborative scientists.

ACKNOWLEDGEMENTS

Addgene would like to thank all of its depositors, requestors, tech transfer officers and partners. We would not be accomplishing our mission without your input, contributions and support. Thank you for 10 years of plasmid sharing and community building. Thank you to all of the Addgenies who do the work every day.

FUNDING

Funding for open access charge: Addgene, Inc.
Conflict of interest statement. J.K. is an employee of Addgene, Inc.

REFERENCES

- Lederberg, J. (1952) Cell genetics and hereditary symbiosis. *Physiol. Rev.*, **32**, 403–430.
- Chandras, C.T., Weaver, M., Zouberakis, D., Smedley, K., Schughart, N., Rosenthal, N., Hancock, J.M., Kollias, G., Schofield, P.N. and Aidinis, V. (2009) Models for financial sustainability of biological databases and resources. *Database*, **2009**, bap017.
- Parsons, J.P. and Duke, C.S. (2013) Strategies for developing and innovating living stocks collections. *Bull. Ecol. Soc. Am.*, **94**, 118–130.
- Baker, M. (2014) Repositories share key research tools. *Nature*, **505**, 272.
- Herscovitch, M., Perkins, E., Baltus, A. and Fan, M. (2012) Addgene provides an open forum for plasmid sharing. *Nat. Biotechnol.*, **30**, 316–317.
- Seiler, C.Y., Park, J.G., Sharma, A., Hunger, P., Surapaneni, P., Sedillo, C., Field, J., Algar, R., Price, A., Steel, J. *et al.* (2014) DNASU plasmid and PSI:Biological-Materials repositories: resources to accelerate biological research. *Nucleic Acids Res.*, **42**, D1253–D1260.
- Crouzier, T. (2014) Researchers in the cloud. *Trends Biochem. Sci.*, **39**, 344–346.
- Fan, M. (2014) *The 10 Most Distributed Plasmid Technologies in Addgene's First 10 Years*. <http://blog.addgene.org/top-10-the-most-distributed-plasmid-technologies-in-addgenes-first-10-years> (15 August 2014, date last accessed).
- Editorial (2013) Receptive to replication. *Nat. Biotechnol.*, **31**, 943.
- Collins, F.S. and Tabak, L.A. (2014) NIH plans to enhance reproducibility. *Nature*, **505**, 612–613.
- Prinz, F., Schlange, T. and Asadullah, K. (2011) Believe it or not: how much can we rely on published data on potential drug targets. *Nat. Rev. Drug Discov.*, **10**, 712.
- Kamens, J. (2014) Top 25 plasmids of all time-Addgene statistics reveal research trends. <http://blog.f1000research.com/2014/03/28/top-25-plasmids-of-all-time-addgene-statistics-reveal-research-trends/> (15 August 2014, date last accessed).
- Furman, J.L. and Stern, S. (2011) Climbing atop the shoulders of giants: the impact of institutions on cumulative research. *Am. Econ. Rev.*, **101**, 1933–1963.
- Kamens, J. (2014) Addgene: Making Materials Sharing “Science-As-Usual”, *PLOS Biol.*, in press.
- Wilkenson, R. and Wiedenheft, B. (2014) A CRISPR method for genome engineering. *F1000Prime Rep.*, **6**, doi:10.12703/P6-3.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
- Hsu, P.D., Lander, E.S. and Zhang, F. (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, **157**, 1262–1278.