Signature-tagged mutagenesis: barcoding mutants for genome-wide screens

Piotr Mazurkiewicz*, Christoph M. Tang*, Charles Boone‡ and David W. Holden*

Abstract | DNA signature tags (molecular barcodes) facilitate functional screens by identifying mutants in mixed populations that have a reduced or increased adaptation to a particular environment. Many innovative adaptations and refinements in the technology have been described since its original use with *Salmonella*; they have yielded a wealth of information on a broad range of biological processes — mainly in bacteria, but also in yeast and other fungi, viruses, parasites and, most recently, in mammalian cells. By combining whole-genome microarrays and comprehensive ordered libraries of mutants, high-throughput functional screens can now be achieved on a genomic scale.

Genetic footprinting
A technique in which a large population of cells is subjected to transposon mutagenesis, then divided and subjected to different selective pressures. PCR on individual genes determines the presence or absence of corresponding mutant strains after selection.

Virulence
The degree of pathogenicity, or ability to cause a disease.

**Department of Infectious Diseases, Centre for Molecular Microbiology and Infection, Imperial College London, Armstrong Road, London SW7 2AZ, UK.
‡Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ontario M5G 1L6, Canada.
Correspondence to D.W.H.
e-mail:
d.holden@imperial.ac.uk*
doi:10.1038/nrg1984

A striking and consistent finding to emerge from the genome-sequencing projects is that the function of most genes cannot be determined from analysis of the primary sequence alone. Instead, clues can be obtained from a range of other approaches, of which the most informative is usually the identification of a mutant phenotype. Phenotypic analysis of mutants that have been obtained by either forward or reverse genetics must therefore continue to have a central role in the post-genome-sequencing, functional genomics era. Meeting this requirement is far from trivial. For even the intensively studied model organism *Escherichia coli* K12, over 50% of its ORFs remain uncharacterized (see The Institute for Genomic Research web site), and it is still a major undertaking to analyse a corresponding number of individual mutants that carry single gene deletions for many interesting and biologically relevant phenotypes.

An attractive alternative to analysing mutants individually is to analyse them in pools. However, to achieve this, one needs a means to distinguish between the different mutants. Genetic footprinting[1] was developed as one approach for more efficient identification of mutants in mixed populations. However, it is restricted in that only one gene is analysed at a time.

An alternative to the slow and laborious analysis of individual mutants is provided by signature-tagged mutagenesis (STM), which was originally designed to enable high-throughput, parallel analysis of mutant strains of pathogenic microorganisms[2]. In STM, each mutant is tagged with a different DNA sequence in such a way that all tags can be co-amplified from the DNA of mixed populations of mutants in a single PCR. They can also be simultaneously labelled to provide specific probes for the detection of mutants, before and after they have been subjected to selection[2]. Therefore, the sequence tag acts as a molecular barcode to monitor the presence of each mutant in the mixed population.

In the original description of the method[2], the tags consisted of short DNA segments containing a 40 bp variable central tag that was flanked by invariant 'arms' of 20 bp in length, which enable the co-amplification and labelling of the central portions by PCR. The junctions of the variable and invariant regions were marked by restriction sites that could be used to release the arms from the central regions following amplification and labelling. These two features allow tag-specific probes to be generated (FIG. 1a). Although the majority of sequences that were generated in this way produced efficiently labelled tags that did not cross-hybridize with each other, this was not true for all sequences, and a pre-screening process was used to remove mutants that carried tags that did not amplify or label efficiently.

The feasibility of the STM method was evaluated using the mouse model of infection by *Salmonella enterica,* because previous research had shown that systemic infection of mice can result from the proliferation of a significant proportion of the bacteria that comprise the inoculum[3]. The process that was used to identify a large number of *S. enterica* virulence genes is illustrated in FIG. 1b.

STM has subsequently been used in many screens to provide functional information on thousands of genes, in particular from pathogenic bacteria and the yeast *Saccharomyces cerevisiae*. Here we describe the refinements that have been made to the methodology, particularly with respect to the use of different mutagens, signature tags and detection methods (for an overview, see TABLE 1). We review the broad range of viral, prokaryotic and eukaryotic systems to which tagging mutagenesis has been applied, and highlight its recent use in conjunction with high-throughput RNAi screens. Finally, we consider what the future might hold for the application of this technology in high-throughput genetic screens.

Technical adaptations

Methods of mutagenesis. Many approaches to random or directed insertional mutagenesis are available. Of course, the choice of the mutagenesis method will depend on its applicability to the organism that is under investigation. Those that have been used in combination with STM include *in vivo* and *in vitro* transposon mutagenesis (in this case it is essential that only one insertion occurs in each genome), shuttle mutagenesis, insertion–duplication mutagenesis by homologous recombination, gene replacement by homologous recombination and illegitimate or non-homologous recombination (FIG. 2). The chosen method should enable the efficient generation of large libraries of different mutants.
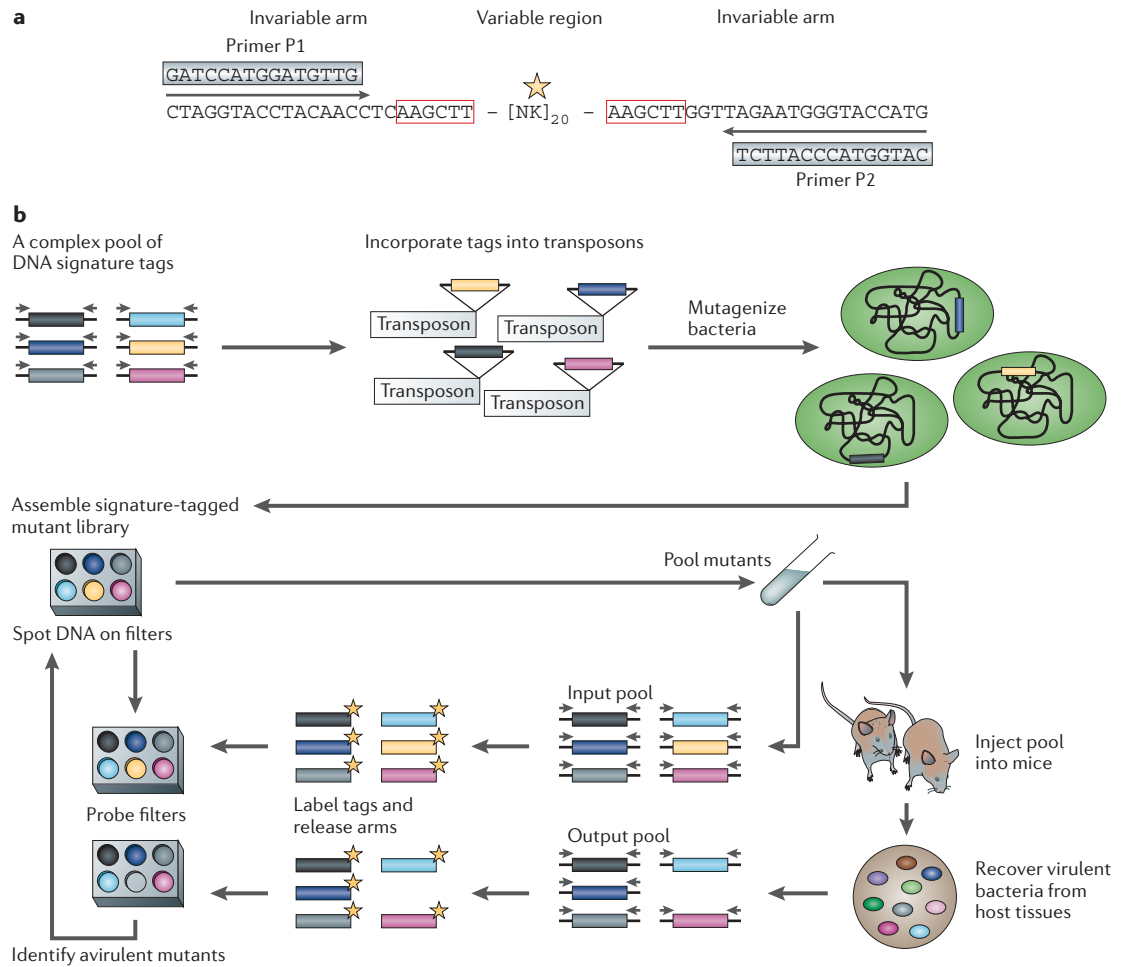

Figure 1 | **Original signature-tagged mutagenesis of *Salmonella*. a** | Design of a signature tag. Each tag has a unique central sequence of 40 bp ([NK]$_{20}$; N = A, C, G, or T; K = G or T), flanked by invariable arms of 20 bp, which are common to all the tags. These arms allow the sequence tags to be amplified and labelled with radioactive nucleotides (marked with a star) by PCR with primers P1 and P2. Following labelling and before hybridization, the invariant arms are removed by digestion with a restriction enzyme that recognizes sequences (shown in red boxes) between the variable region and the invariable arms. **b** | Signature-tagged mutagenesis screening in mice. A complex pool of tags (shown as coloured rectangles) is ligated to transposons. The tagged transposons are then used to mutagenize bacteria, which are subsequently assembled into a library. Only bacteria with tags that are efficiently amplified by PCR and are not cross-reactive with other tags in hybridization experiments are selected for inclusion in the pool that is used to infect the mice. Genomic DNA is isolated from this pool (input pool) and from the bacteria that are recovered from the animals (output pool). The tags from these two DNA pools are amplified and radiolabelled to create probes for hybridization. DNA from the colonies of the mutant library that hybridize to the probes from the input pool but not to the probes from the recovered pool represent mutants with attenuated virulence.

Tags and screening. The problem of unreliable amplification and/or labelling of tags can be solved by the empirical selection of tags on the basis of their efficient amplification, labelling and lack of cross-hybridization to other tags. Such pre-selected tags are then used separately to generate a potentially infinite number of mutant strains, which are arrayed according to the tags they carry (FIG. 3Aa). Another advantage of this approach is that because the identity of the tag in each mutant is known, labelled tags can be hybridized to purified plasmids that carry tags[4], or to purified tag DNA[5], rather than to the chromosomal DNA of mutant strains. This greatly increases the sensitivity of the assay and allows the use of non-radioactive detection methods, for example, digoxigenin[4] or biotin[6]. The incorporation of longer tags[7,8] or multiple tags at each mutated locus[6,9] is another adaptation that was designed to increase the reliability of the screening process.

Another modification involves the use of high-density oligonucleotide arrays for hybridization analysis[6,9–12]. In principle, it enables thousands of sequences to be analysed in parallel, but in the case of pathogenic bacteria, the number of mutants that can be screened *in vivo* is sometimes restricted by aspects of host anatomy and immunity; this limitation must usually be investigated in pilot experiments before large-scale screening can be initiated. Therefore, to fully exploit the potentially vast scale-up that is offered by microarrays, pools of DNAs or microorganisms from different hosts might need to be combined before hybridization analysis. However, if mutant microorganisms are being tested in environments outside living hosts, these assays can frequently be scaled up to allow analysis of highly complex pools[9,13].

STM without hybridization. As an alternative to hybridization, PCR products can be analysed directly[5] to indicate the presence or absence of tags (FIG. 3Ab). This method relies on using primers that are specific for each tag. Mutagenesis is carried out with transposons that carry different tags of known sequence, and the DNA that is recovered from virulent mutants is subjected to PCRs in which at least one of the primers is tag-specific. The total number of PCRs that are required for analysis is therefore twice the number of mutants being analysed, and the products are visualized by agarose gel electrophoresis[14]. This simple modification has the great advantage of circumventing the need for hybridization after the PCR step. However, this approach is inherently less quantitative, and a large number of PCR products must be analysed by gel electrophoresis — a problem that was addressed by the introduction of multiplex PCR-based STM[15]. This modification uses a small number of tags with known sequences that have been combined with three different selection markers. The mutants that are recovered are identified by a PCR in which a tag-specific primer is combined with three primers that anneal to the selection markers, yielding three different PCR products. For example, a combination of 24 sequence tags and 3 selection markers allowed a pool of 72 mutants to be analysed in 24 PCRs[15].

Another approach that avoids hybridization, and further reduces the number of PCRs that are required, is polymorphic tag-length transposon mutagenesis (PTTM), which has been applied to group A *Streptococcus*[16] (FIG. 3Ac). In this modification of the method, specificity is conferred by the different lengths

Table 1 | **Technical modifications of signature-tagged mutagenesis***

Procedure modified	Examples	Organisms
For efficient mutagenesis		
Transposition	*In vivo* transposition[2,76,82–86]	*Salmonella typhimurium, Mycobacterium tuberculosis, Vibrio cholerae, Yersinia enterocolitica, Legionella pneumophila, Brucella suis, Escherichia coli*
	In vitro transposition[28,87,88]	*Streptococcus pneumoniae, Neisseria meningitidis, Helicobacter pylori*
Mutagenesis without transposons	Shuttle mutagenesis[89]	*N. meningitidis*
	Insertion–duplication mutagenesis[5,90]	*S. pneumoniae*
	Gene-replacement by homologous recombination[10,11]	*Saccharomyces cerevisiae*
	Non-homologous (illegitimate) recombination[53,57]	*Candida glabrata, Cryptococcus neoformans*
For efficient detection of tags		
Tag design	Longer tags[7,8]	*Yersinia pestis, S. typhimurium,*
	Multiple tags[6,9]	*Yersinia pseudotuberculosis, S. cerevisiae*
	Pre-selected pools of tags[4,14]	*Staphylococcus aureus, Pseudomonas aeruginosa*
Detection method	Hybridization to purified plasmids carrying tags[4]	*S. aureus*
	Hybridization to PCR-amplified tags[5]	*S. pneumoniae*
	Non-radioactive probes[4,6]	*S. aureus, Y. pseudotuberculosis*
	Direct detection of PCR products[14,16,89]	*P. aeruginosa*, group A *Streptococcus, N. meningitidis*
	Microarray assembly of tags[9–12]	*S. cerevisiae, Desulfovibrio desulfuricans, Shewanella oneidensis*

*The choice of a specific method of mutagenesis will depend on its applicability to the organism of interest.

of the tags, each of which can be distinguished by the separation of PCR products on acrylamide gels. Only two PCRs (for input and output) are required for each screen.

TraSH, MATT and DeADMAn. Transposon site hybridization (TraSH)[17,18], microarray tracking of transposon mutants (MATT)[19], and designer arrays for defined mutant analysis (DeADMAn)[20] are variations

a Direct mutagenesis; in vivo transposition

Construct a pool of tagged transposons

Introduce tagged transposons into the organism of interest by conjugation or transformation

Select clones with integrated tagged transposons

b Shuttle mutagenesis; in vivo transposition

Create a DNA library of the organism of interest in a shuttle vector

Introduce the library into the organism in which the transposon mutagensis will be performed

Introduce tagged transposons into the organism that carries the library

Select clones with integrated tagged transposons

Introduce the mutagenized library into the organism of interest

Chromosome

Mutants are created by homologous recombination

c In vitro transposition

Isolate genomic DNA or DNA library from the organism of interest

Mix with a pool of tagged transposons in vitro

The isolated genomic DNA or DNA library becomes a target for transposon insertions

Introduce the mutagenized DNA into the organism of interest

Chromosome

Mutants are created by homologous recombination

d Insertion–duplication mutagenesis

Construct a pool of tagged suicidal plasmids

Prepare fragments of genomic DNA from the organism of interest (random or PCR-amplified specific fragments)

Clone fragments of genomic DNA into tagged plasmids

Introduce the tagged DNA library into the organism of interest

Chromosome

Mutants are created by homologous recombination

e PCR-mediated directed gene disruption

Construct primers containing sequences that are homologous to the selection marker, tag sequence and short regions of sequence that are identical to those upstream and downstream of the targeted gene

Downstream primer   Upstream primer

Selection marker

Use PCR to introduce unique tags and target gene sequences on both ends of the selection marker

Introduce modified selection marker into the target organism

Chromosome

Mutants are created by homologous recombination
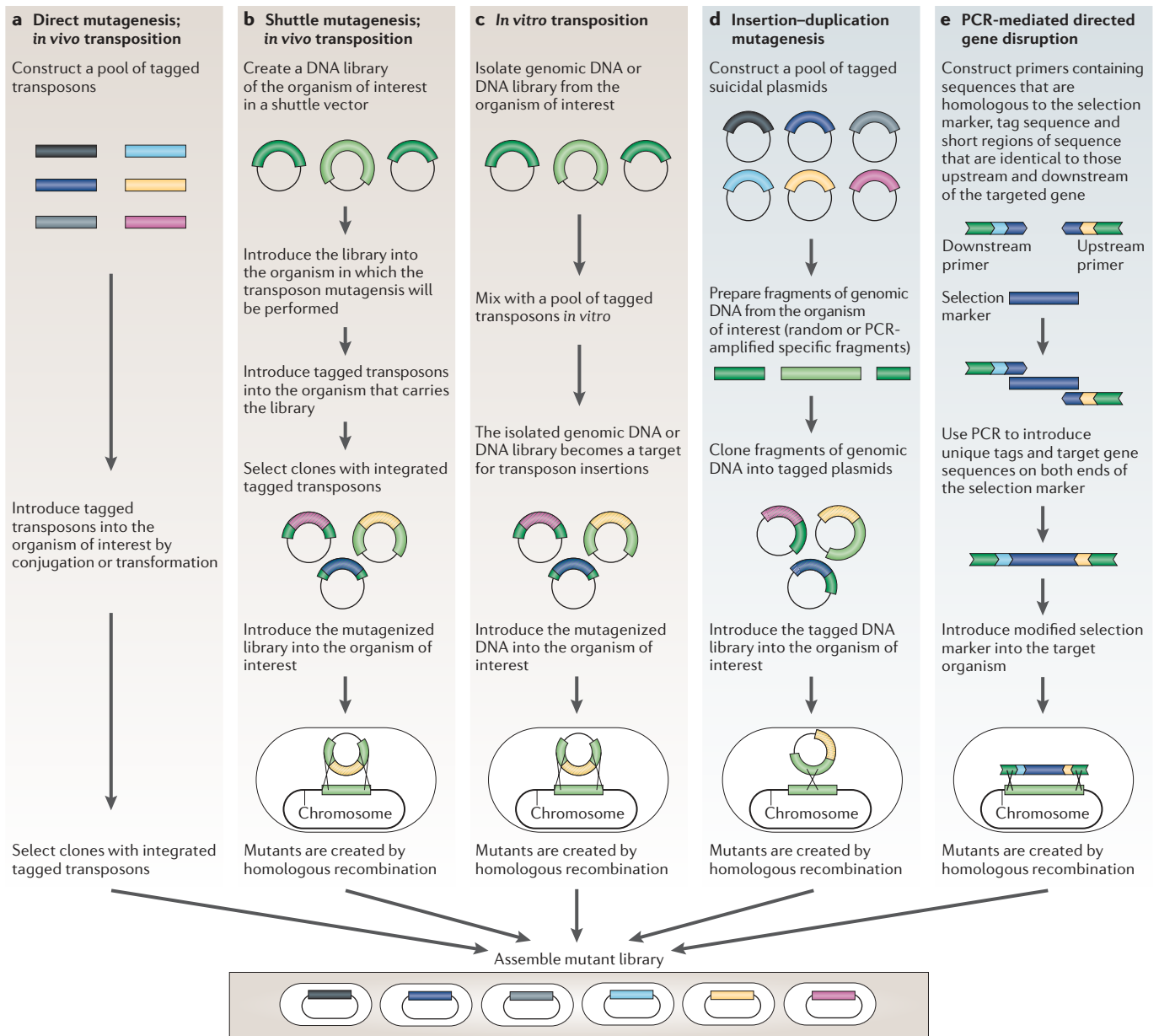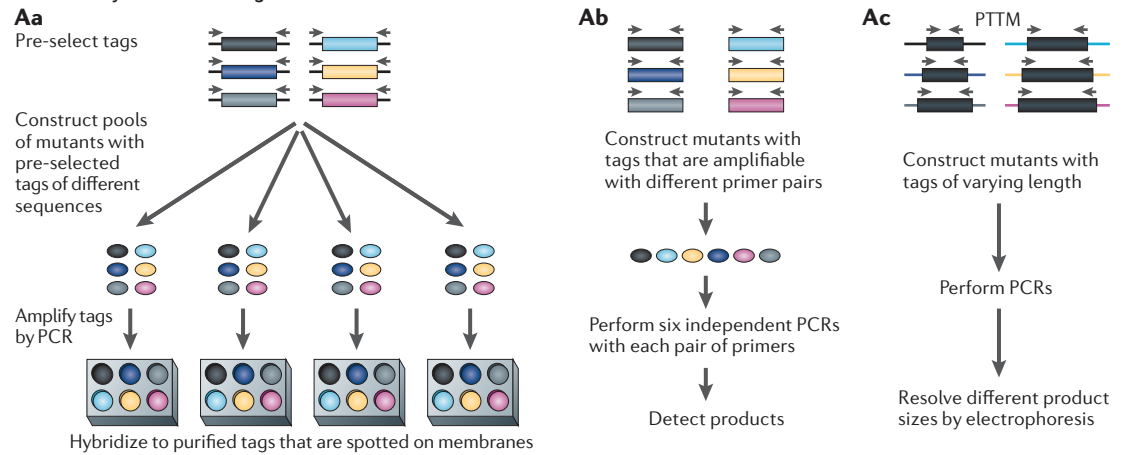
Assemble mutant library

Figure 2 | **Methods for generating pools of tagged mutants.** Panels **a**, **b** and **c** show methods that involve random transposition. The methods in panels **d** and **e** use homologous recombination. DNA tags are represented by different coloured segments, whereas the targeted sequences are shown in green. Tags can be introduced into the genome by direct *in vivo* transposition of the target organism (panel **a**). Alternatively, transposition can be carried out on the target DNA library *in vivo*, in an organism for which an efficient transposition method exists (panel **b**), or *in vitro*, on isolated DNA (panel **c**). In the methods depicted in panels **b** and **c**, the mutagenized DNA is subsequently reintroduced into the target organism to allow the incorporation of tags into the chromosome by homologous recombination. Insertion–duplication mutagenesis (panel **d**) involves ligating small fragments (random or specific) of target DNA into a pool of tagged plasmids. The plasmids are then introduced into the microorganism of interest, where they integrate into the genome. In PCR-mediated gene disruption (panel **e**), PCR is carried out to amplify a selectable marker, which is flanked with short sequences (~50 bp) that are identical to those immediately downstream and upstream of the targeted gene. When introduced into a cell, the resulting PCR product that incorporates the marker can replace the targeted gene by homologous recombination. For more information on techniques of transposon-based mutagenesis see REF. 81.

on STM that incorporate microarray technology (FIG. 3B). In each case, DNA is extracted from bacterial transposon-mutagenized pools before and after a selective process, and unique sequences that are physically linked to each mutation are amplified and labelled before hybridization to a genomic microarray. By comparing the signal intensity that is generated by probes that have been derived from the mutants pre- and post-selection, those with a selective disadvantage are identified. The main difference between these techniques is the way in which specific probes are generated for each mutation. In the case of TraSH, the genomic DNA that is isolated from the mutant pool is partially digested with a restriction enzyme that makes frequent cuts in the genome. Double-stranded adaptors are then ligated to the ends of the digested DNA. A PCR is carried out with primers that anneal to the adaptors to amplify the DNA regions that flank the transposon insertion. Next, the PCR

A Use of synthetic DNA tags

Aa
Pre-select tags

Construct pools of mutants with pre-selected tags of different sequences

Amplify tags by PCR

Hybridize to purified tags that are spotted on membranes

Ab
Construct mutants with tags that are amplifiable with different primer pairs

Perform six independent PCRs with each pair of primers

Detect products

Ac PTTM
Construct mutants with tags of varying length

Perform PCRs

Resolve different product sizes by electrophoresis

B Use of flanking sequences as tags

Ba TraSH
Digest then ligate linkers (▶)

Perform PCR to generate probes

Hybridize to genomic oligonucleotide arrays

Bb MATT
Perform two rounds of PCR, with the first round including a primer of degenerate sequence

Hybridize to genomic oligonucleotide arrays

Bc DeADMAn
Isolate insertion sites from all mutants

Construct an array containing the insertion sites

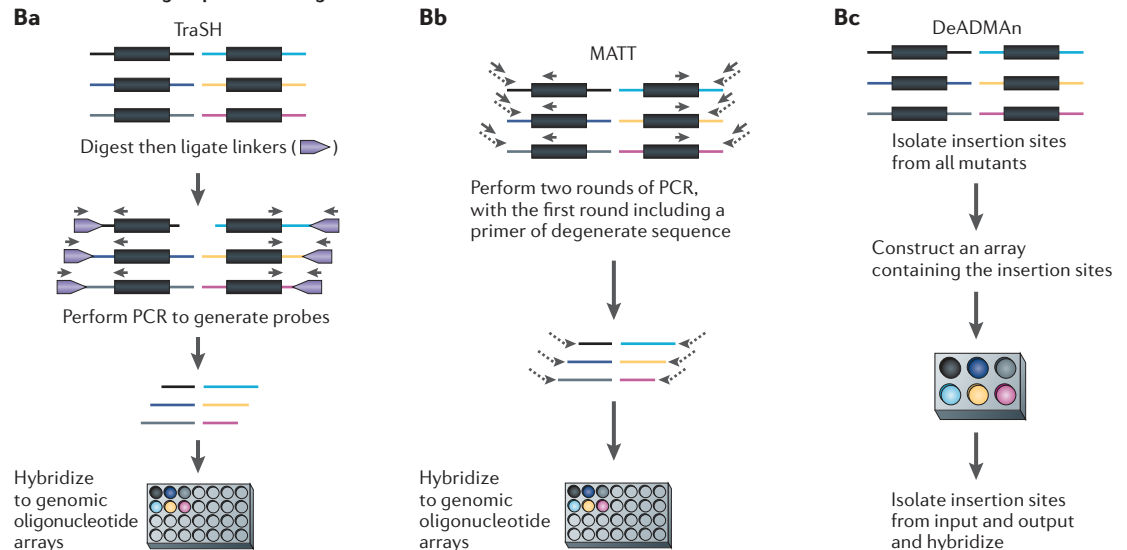Isolate insertion sites from input and output and hybridize

Figure 3 | **Methods for the detection of signature tags.** Several techniques have been designed that incorporate synthetic DNA tags (**A**) or that take advantage of flanking sequences (**B**). **A** | Tags that are efficiently and specifically amplified and labelled can be pre-selected and used repeatedly to generate separate pools of mutants (coloured ovals in part **Aa**). Membranes can then be constructed with purified tags or the plasmids that harbour them. The detection of tags can be carried out without the need for hybridization. Tags can be amplified in multiple PCRs, each containing a different primer pair for a specific tag (**Ab**), or in polymorphic tag-length transposon mutagenesis (PTTM, panel **Ac**), tags of different length are amplified with a single primer pair, giving rise to products of various sizes. **B** | Probes that are generated from the flanking sequences can be used to hybridize to genomic microarrays. In transposon site hybridization (TraSH; panel **Ba**), flanking sequences are amplified by ligating linkers to digested genomic DNA from pools of mutants. In microarray tracking of transposon mutants (MATT; panel **Bb**), flanking sequences are amplified by arbitrary PCR (which involves two rounds of PCRs, with the first round including a primer of degenerate sequence (dashed arrow) and a transposon specific primer (solid arrow)). In designer arrays for defined mutant analysis (DeADMAn; panel **Bc**), the sequences that flank each mutation are isolated and assembled onto an array, which is then used for subsequent hybridizations.

Box 1 | Examples of biological insights from STM screens

Salmonella

Signature-tagged mutagenesis (STM) of *Salmonella typhimurium* led to the discovery of a specialized type 3 secretion (T3S) system, which is encoded by a horizontally acquired pathogenicity island called SPI2. The T3S system transfers over 15 virulence proteins across the vacuolar membrane that encloses intracellular bacteria. At least some of the virulence proteins are involved in controlling vacuolar membrane dynamics through regulation of molecular motors[74], and together they enable intracellular bacterial replication during infection. A strain of *Salmonella typhi* that lack the SPI2 T3S system shows considerable promise as a new live attenuated vaccine for typhoid fever[75].

Mycobacterium tuberculosis

The lung is the predominant organ that is affected by *M. tuberculosis*, and pulmonary infection causes around two million deaths each year (see the World Health Organisation web site). Two separate STM studies demonstrated that a complex cell wall lipid, phthiocerol dimycocerosate (DIM), is necessary for survival of the bacterium in the lung, but not in other tissues such as the spleen[76,77]. Mutations that attenuate virulence were found to cluster in a locus that is necessary for the biosynthesis and export of this lipid, which is produced mainly by pathogenic members of the *Mycobacterium* genus. One of the affected genes (*lppX*) encodes the first member of a new family of lipoproteins that function as carriers of lipophilic molecules across the mycobacterial cell envelope[78]. DIM production evidently protects *M. tuberculosis* from the toxic effects of nitric oxide that are produced by macrophages[79].

Shigella flexneri

Shigella flexneri is the leading cause of bacillary dysentery worldwide, and causes disease after invading epithelial cells. The invasion is dependent on the activity of a T3S system that delivers cell bacterial effectors into the host, eliciting dramatic rearrangements of the cytoskeleton. Detailed analysis of colonization-defective mutants identified those with minor alterations in the composition of the outer-membrane molecule lipopolysaccharide (LPS)[80]. The work revealed an intimate relationship between the ability of the LPS molecule to protect the pathogen from innate immune responses, and to allow correct presentation of the T3S system at the bacterial surface. The balance is achieved through conformational changes in the LPS molecule, which are induced by glucosylation[80].

Saccharomyces cerevisiae

The large number of comprehensive screens in budding yeast have identified numerous genes that are involved in various processes. For example, a collection of heterozygous yeast deletion strains were examined to identify those with defective growth in complete medium[38]. This was undertaken to identify the extent of haploinsufficiency, the state in which the loss of one of two alleles leads to a phenotype. Approximately 3% of the strains showed haploinsufficiency; unsurprisingly, over half of these mutants had defects in essential genes. The others were all defective for growth as homozygous mutants and were predominantly affected in metabolic functions, particularly those involving multiprotein complexes. Haploinsufficiency of many strains could be relieved by simply reducing their growth rate by propagation in minimal medium.

Pathogenicity island
A distinct region of the genome, usually with an altered G and C content and containing one or more virulence genes, that was probably acquired by horizontal gene transfer.

Cell envelope
The cell membrane and outer membrane, if one is present, together with the cell wall; Mycobacteria have a very complex cell wall that is rich in glycolipids, especially mycolic acids.

products are used as templates for transcription by T7 RNA polymerase, which transcribes from the transposon into the genomic DNA that flanks the transposon insertion. Finally, labelled cDNA is generated by reverse transcriptase PCR (RT-PCR) and hybridized to DNA microarrays.

This method has been simplified by eliminating the adaptors and using degenerate primers for probe generation by RT-PCR. The microarray analysis was also optimized by labelling input and output probes with different fluorescent dyes to allow their simultaneous hybridization to the DNA array[21].

The distinctive feature of MATT[19] is that the DNA that lies adjacent to a transposon insertion is amplified by a two-step PCR, without the RT-PCR step that is carried out in TraSH. In the first reaction, genomic DNA is amplified by a transposon-specific primer and a primer containing a degenerate 3′ region and an invariant 5′ anchor region. The products of this reaction are used as a template for a second PCR, which uses primers that are complementary to an amplified region of the transposon and the conserved portion of the degenerate primer. The amplified DNA from this reaction is coupled to a dye and hybridized to a microarray. This method of obtaining the flanking sequences is known as arbitrary PCR, because of the degenerate nature of the primer that is used in the first round of amplification.

In the more laborious DeADMAn approach, the sequences that flank transposon insertion sites are determined for every mutant in the pool, and oligonucleotides that are based on these flanking sequences are used to build a microarray. DNA from pools of mutants is extracted in the normal way, cut with restrictions enzymes and ligated to adaptors. Nested PCR primers are then used to amplify the DNA that flanks the transposon, and the products are labelled with fluorescent dyes before hybridization to the microarray[20]. These approaches all have an important advantage in that they circumvent the need to synthesize tags for each transposon.

Choice of tagging and detection methods. The principles of tag design and detection remain universal and can, in principle, be applied to any organism. The original tag design[2] has proved remarkably robust and has been used in the majority of bacterial STM studies[22]. It has also been adapted for use in yeast[9,10] and in mammalian RNAi screens[23,24] (see below). PTTM, TraSH, MATT and DeADMAn have not been used extensively outside the laboratories in which they were developed.

With respect to the detection of tags, hybridization-based methods are the most popular. However, to our knowledge, there has been no comparative study of the sensitivity and ease of different tagging and detection methods with the same model organism and selection procedure. For this reason, the choice of a specific method seems to be largely based on personal preference and expertise.

Application to pathogenic bacteria

So far, STM has found its broadest applications in studying pathogenic bacteria. Numerous screens have been carried out involving all the leading human pathogens that are genetically tractable, which has resulted in the identification of over 2,000 virulence and colonization determinants (see REFS 25,26 for reviews). Some of these studies are listed in TABLE 1. Examples of important biological insights that have emerged from follow-up studies are highlighted in BOX 1. Bacterial pathogens frequently colonize more than one cell type, tissue or organ of a host during the course of infection, and different virulence factors are needed to enable growth and survival in these different environments. To identify genes that are involved in these processes, different models of infection can be used to screen the same mutant library in the same host by STM, in order to identify tissue- or organ-specific virulence factors[27,28] (BOX 1).

Some bacteria show remarkable host adaptation and can only cause disease in certain species, whereas others can cause disease in a range of hosts. The genetic basis for this has received little attention and remains largely unexplained. To identify factors that mediate host adaptation, the same pools of mutants can be inoculated into different host species to reveal mutants that are attenuated for virulence in only one host[29,30].

STM can also be used in knockout mutant mice to identify genes that counter innate immune effectors. In this technique, called 'differential STM' (FIG. 4), pools of mutant pathogenic bacteria are used to infect different immunodeficient mouse strains, and mutant bacterial strains are identified on the basis of their ability to proliferate in the tissues of mice of one genetic background, but not another[31,32]. Subsequent analysis of the specific functions of genes that are affected in the so-called 'counter-immune' mutant mice can be enhanced by knowledge of the function of the relevant host genes.

Application to yeast

Although STM was originally developed for use with microbial pathogens, it has also been exploited to assist large-scale functional genomic studies of the budding yeast *S. cerevisiae*. In *S. cerevisiae*, targeted mutagenesis is achieved by taking advantage of the efficient homologous recombination in this organism[10]. This has facilitated the construction of a nearly complete collection of deletion mutants, covering 96% of ORFs in the yeast genome[9,13]. To make the mutant library, short regions of DNA that corresponded to the ends of yeast genes were used in PCRs to amplify a selectable marker. Two unique signature tags were also incorporated into each construct, and a total of almost 6,000 deletion strains have been created by an international consortium[13]. Used in conjunction with a set of DNA microarrays that represent all the signature tags in the library, every member of this collection of mutants can be simultaneously tested for altered growth rates in response to any environmental condition. Those that have been tested to date have identified thousands of genes that are involved in processes as diverse as nutrient utilization, responses to osmotic stress, high pH, high osmolarity and ion stress[13], resistance to DNA-damaging agents[33–36], drug resistance, sporulation and post-germination growth[37], mechanisms of haploinsufficiency[38], the non-homologous end-joining pathway[39], the response to proteasome inhibition[40], endoplasmic reticulum biosynthesis[41], and mitochondrial respiration[42] (BOX 1).

The yeast consortium constructed a library of both homozygous and heterozygous signature-tagged diploid mutants[9,13]. An interesting aspect of heterozygous deletion mutants is that some are sensitized to drugs that inhibit the product of the heterozygous gene[11]. This gene dosage, or 'haploinsufficiency' effect, can be used to identify genes with products that represent potential drug targets, a process that is referred to as 'chemogenomic profiling'. Two large-scale screens have been conducted to identify the targets of various clinically or agriculturally relevant compounds with distinct chemical structures[43,44]. Although haploinsufficiency or chemogenomic profiling has its limitations, this approach is
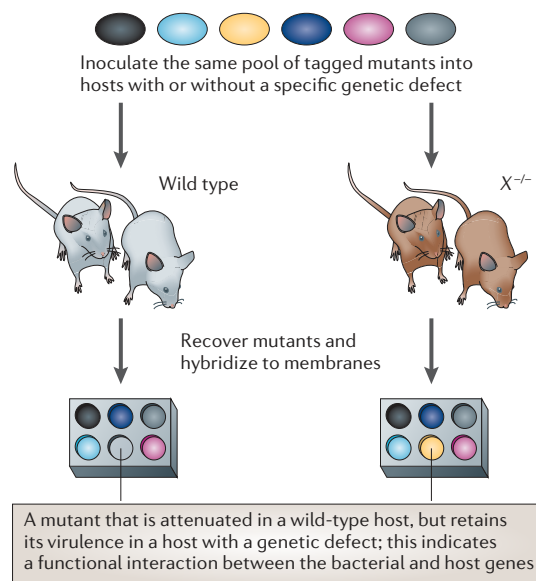

Figure 4 | Differential STM screen. The same set of mutants is inoculated into wild-type animals, and those with a specific genetic defect ($X^{-/-}$). Counter-immune mutants are attenuated in a wild-type host, but retain their virulence in the knockout animals. This indicates that there is an interaction between the product of the gene that is affected in the counter-immune mutant and the function that is lost in the knockout animal.

likely to facilitate rapid identification of many targets for new small-molecule inhibitors.

The set of yeast deletion mutants (haploids or homozygous diploids) can also be examined for chemical sensitivity using a pooled mutant set and a barcode microarray readout. This approach identifies genes with products that buffer the cells from the toxic effects of the compound, thereby generating a chemical genetic profile. Compounds with similar chemical genetic profiles often target the same pathway and therefore have a similar mode of action, whereas deletion mutants that show a similar signature of chemical sensitivities often have similar cellular functions[45–47].

Synthetic lethality is a useful genetic tool that has been pioneered in yeast[48]. It can result from a lack of function in parallel biochemical pathways or redundant components of the same essential pathway, and provides an indication that two genes are functionally related. The synthetic genetic array (SGA) system enables large-scale analysis of synthetic lethal phenotypes using high-density yeast arrays[49,50]. The availability of the signature-tagged mutant library and DNA microarrays to analyse the relative abundance of tags has also enabled synthetic lethal screens to be carried out on a genome-wide scale, a procedure that is referred to as synthetic lethality analysis by microarray (SLAM)[51]. An adapted version, dSLAM (diploid-based SLAM), involves transformation of heterozygous diploids carrying an SGA reporter, which enables selection for haploid double mutants after sporulation of transformed diploids, thereby improving the robustness of this method[52].

Innate immunity
'General' or 'non-specific' immune responses; the first line of defence against a pathogen.
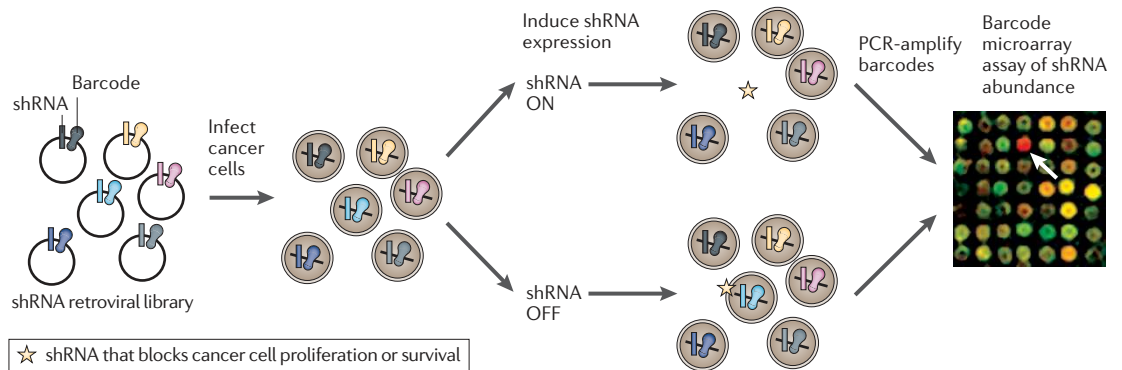
Figure 5 | **RNAi screen with barcoding.** Retroviral vectors that encode a library of short hairpin RNAs (shRNAs) that are under the control of an inducible promoter are introduced into a cancer cell line. The cells are then divided in two subpopulations: one is subjected to induction of shRNA expression and the other is used as a control cell population. An shRNA that reduces the expression of a protein that is critical for proliferation or survival of the cancer cells will be eliminated from the induced, shRNA-expressing culture. Genomic DNA is isolated from the two populations at different time points, and PCR is used to amplify the barcodes that are present in the genomic DNA. Amplified DNAs from the induced and control cultures are labelled with different fluorescent dyes and cohybridized to barcode oligonucleotides on microarrays to determine the relative abundance of each barcode in the two populations. This indicates the relative depletion or enrichment of cells that express a given shRNA[65].

Application to other organisms

A modified STM screen was used to find adhesins of the haploid, opportunistically pathogenic yeast *Candida glabrata*[53]. In this screen, different DNA tags were introduced into a dispensable chromosomal locus to provide 96 tagged strains. They were subsequently mutagenized by non-homologous integration of a vector that carried a selectable marker into the chromosome. Screening of 4,800 mutants yielded 31 that had altered adherence to human epithelial cells, and led to the discovery of a novel family of adhesive surface glycoproteins[54].

Tools that are based on either homologous recombination or illegitimate recombination between plasmid and chromosomal DNA have also been developed for insertional mutagenesis-based STM screens on three other pathogenic fungi: *Candida albicans*[55], *Aspergillus fumigatus*[56] and *Cryptococcus neoformans*[57].

The ability to clone entire viral genomes on BACs means that STM can be used to analyse gene function in DNA viruses with large genomes. The application of STM to the mouse γ-herpesvirus 68 has allowed screens to be carried out in cultured fibroblasts and the mouse host[58,59], and might in the future facilitate the analysis of genes that are involved in reactivation from latency — an important aspect of the infectious process of herpesviruses.

Protozoan parasites with haploid stages in their life cycles can, in principle, be suitable for STM screening, as long as efficient methods of insertional mutagenesis exist. Proof of principle for STM screening of *Toxoplasma gondii* (a zoonotic pathogen and a model for apicomplexan parasites, which include the genus *Plasmodium*) has already been established by insertional mutagenesis with an STM plasmid[60].

Barcoding for RNAi screens

Until recently, it was difficult to conduct large-scale genetic screens in diploid mammalian cells because of the relatively large size of their genomes and the need to inactivate both alleles of a gene. However, the advent of transient gene silencing using small interfering RNAs (siRNAs) has, to a significant degree, overcome these problems[61]. In these screens, either chemically synthesized dsRNA or a retrovirus-based vector (from which short hairpin RNAs (shRNAs) are transcribed) are introduced into mammalian cells. The cellular mRNA that contains the sequence of the dsRNA or shRNA is usually recognized and degraded, resulting in efficient knockdown of gene expression. Libraries of thousands of siRNAs can be assembled for analysis in 96-well plates, but subsequent infection or transfection and analysis of mammalian cells is a labour-intensive and time-consuming process. Two groups have demonstrated the feasibility of using signature tagging in conjunction with siRNA to facilitate high-throughput screening. Berns *et al.*[23] used the unique sequence of each shRNA as a signature tag, whereas Paddison *et al.*[24] incorporated unique synthetic 60-bp tags into each shRNA vector.

The utility of this approach has been highlighted by its application to the discovery of a novel human tumour suppressor[62], and a new target of the antiproliferative drug nutlin-3 (REF. 63). Second-generation libraries comprising over 140,000 shRNA expression plasmids, each carrying a randomly generated 60-bp signature tag, are now available[64]. Whereas these two studies identified shRNAs that are selectively enriched under certain experimental conditions, recent work by Staudt and colleagues[65] describes a method for identifying shRNAs that are depleted during the course of an experiment. These authors used a library of barcoded vectors encoding shRNAs under the control of an inducible promoter. The library is introduced into a cell line, which is then divided into two groups: one is grown in the presence and the other in the absence of the shRNA expression inducer. Barcode abundances, which are assayed using

---

Adhesins
Molecules that mediate the attachment of an organism to a surface.

Latency
The persistence of a pathogen in the host cell without necessarily causing signs of disease.

Zoonotic
A pathogen that primarily affects animals but is transmissible to humans.

Apicomplexan
A type of protozoan parasite that uses a unique form of actin-based motility and has a complex life cycle.

Small interfering RNAs
Small antisense RNAs (20–25 nucleotides) that are generated from specific dsRNAs that trigger RNAi. They serve as guides for the cleavage of homologous mRNA in the RNA-induced silencing complex.

Short hairpin RNAs
Small RNAs that form hairpins that can induce sequence-specific silencing in mammalian cells through RNAi, when produced exogenously and transfected into the cell or when expressed endogenously.

microarrays, give an indication of the fitness of cells expressing specific shRNAs (FIG. 5). This 'loss-of-function' RNAi screen (like conventional STM screens) relies on negative selection, and could be used to discover new oncogenic pathways that promote tumour malignancy[65]. These recently established resources should greatly facilitate large-scale analysis of mammalian gene function in the future.

Perspectives

Future STM screens of organisms other than yeast will be greatly helped by the development of comprehensive, ordered libraries of mutants and whole-genome DNA microarrays. Some fully sequenced bacterial genomes have already been used to construct ordered mutant libraries[66], and even if these are not made with chemically synthesized tags, STM can be carried out using the flanking sequences at insertion sites to provide tags. Recently, an ordered, non-redundant transposon mutant library of 4,596 predicted ORFs of *Pseudomonas aeruginosa* has been constructed (corresponding to 77% of all predicted genes) in a way that allows TraSH analysis to be carried out[67]. There is, in principle, no reason that similar libraries could not be constructed for other important bacterial species; such libraries could also be used for screening synthetic lethal mutations, similar to the yeast SLAM technique[51].

The creation of large collections of mouse mutants, generated either by targeted[68] or random[69] mutagenesis, offers tremendous opportunities to study pathogenicity genes that target specific aspects of host immunity, and to discover novel mechanisms of resistance to microbial infection. For example, eight mutations that cause susceptibility to mouse cytomegalovirus (MCMV) infection have been identified in a screen of 3,500 mice[70]. Further screening is expected to define the MCMV 'resistome': the total number of genes with non-redundant functions in resistance to this pathogen. We can anticipate that the resistomes to other pathogens of mice will be characterized in the future. The corresponding mutant mouse strains could then be used in conjunction with ordered collections of signature-tagged mutant pathogens to provide high-throughput counter-immune screens for the identification of pathogen virulence factors that target specific immune functions.

One potential application of tagged strains has been largely neglected — the study of the populations of genetically identical strains. By varying the pool complexity and inoculum dose, it might be possible to exploit the tags to obtain information on anatomical and immunological bottlenecks, as well as pathogen population dynamics and transmission during infection[71].

The integration of barcoding with RNAi screens is an emerging technology that promises to deliver unprecedented insights into eukaryotic cellular processes. The ability to knockdown up to three genes simultaneously using multi-shRNAs[72,73] provides an opportunity for RNAi-based synthetic-lethality-like studies. Perhaps the most challenging aspect of multi-RNAi barcoding will be the design of efficient screening procedures to deliver specific selective pressures for physiological processes of interest.

Conclusion

STM has proved to be a robust and powerful high-throughput screening technique for the analysis of genes that are not essential for life, but are needed for growth in specific environments. Its application has uncovered unexpected phenotypes for many previously uncharacterized genes, particularly those for which bioinformatics has been essentially uninformative. But, it is important to bear in mind that for every identified gene with a sequence that does not reveal function, a great deal of careful work is required to identify more specific mutant phenotypes, which in turn provide clues to biochemical function. Further research is then necessary to define the precise role of the gene product by determining its cellular location, biochemical activity, interacting partners and structure. Given the amount of work that is involved in analysing gene function, insertional mutagenesis will continue to provide the basis for innumerable important biological questions for decades to come.

1. Smith, V., Botstein, D. & Brown, P. O. Genetic footprinting: a genomic strategy for determining a gene's function given its sequence. *Proc. Natl Acad. Sci. USA* **92**, 6479–6483 (1995).
2. Hensel, M. *et al.* Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269**, 400–403 (1995).
   The first paper to describe the use of STM. Its application to *S. typhimurium* led to the discovery of genes that encode a specialized type 3 secretion system that is necessary for virulence in mice.
3. Mahan, M. J., Slauch, J. M. & Mekalanos, J. J. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science* **259**, 686–688 (1993).
4. Mei, J. M., Nourbakhsh, F., Ford, C. W. & Holden, D. W. Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol. Microbiol.* **26**, 399–407 (1997).
5. Polissi, A. *et al.* Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infect. Immun.* **66**, 5620–5629 (1998).
6. Karlyshev, A. V. *et al.* Application of high-density array-based signature-tagged mutagenesis to discover novel *Yersinia* virulence-associated genes. *Infect. Immun.* **69**, 7810–7819 (2001).
7. Leigh, S. A., Forman, S., Perry, R. D. & Straley, S. C. Unexpected results from the application of signature-tagged mutagenesis to identify *Yersinia pestis* genes required for adherence and invasion. *Microb. Pathog.* **38**, 259–266 (2005).
8. Bearson, S. M., Bearson, B. L. & Rasmussen, M. A. Identification of *Salmonella enterica* serovar Typhimurium genes important for survival in the swine gastric environment. *Appl. Environ. Microbiol.* **72**, 2829–2836 (2006).
9. Winzeler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
10. Shoemaker, D. D., Lashkari, D. A., Morris, D., Mittmann, M. & Davis, R. W. Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nature Genet.* **14**, 450–456 (1996).
11. Giaever, G. *et al.* Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nature Genet.* **21**, 278–283 (1999).
   Signature-tagged heterozygous yeast strains were used to screen for drug sensitivities that were caused by the loss of one allele (haploinsufficiency).
12. Groh, J. L., Luo, Q., Ballard, J. D. & Krumholz, L. R. A method adapting microarray technology for signature-tagged mutagenesis of *Desulfovibrio desulfuricans* G20 and *Shewanella oneidensis* MR-1 in anaerobic sediment survival experiments. *Appl. Environ. Microbiol.* **71**, 7064–7074 (2005).
13. Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002).
   This paper describes a systematic construction of a nearly complete collection of tagged yeast mutants and their use with DNA microarrays to determine the abundance of each deletion strain in mixed populations.
14. Lehoux, D. E., Sanschagrin, F. & Levesque, R. C. Defined oligonucleotide tag pools and PCR screening in signature-tagged mutagenesis of essential genes from bacteria. *Biotechniques* **26**, 473–478, 480 (1999).
15. Potvin, E. *et al.* *In vivo* functional genomics of *Pseudomonas aeruginosa* for high-throughput screening of new virulence factors and antibacterial targets. *Environ. Microbiol.* **5**, 1294–1308 (2003).
16. Hidalgo-Grass, C. *et al.* A locus of group A *Streptococcus* involved in invasive disease and DNA transfer. *Mol. Microbiol.* **46**, 87–99 (2002).

17. Badarinarayana, V. *et al.* Selection analyses of insertional mutants using subgenic-resolution arrays. *Nature Biotechnol.* **19**, 1060–1065 (2001).

18. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Comprehensive identification of conditionally essential genes in mycobacteria. *Proc. Natl Acad. Sci. USA* **98**, 12712–12717 (2001).

19. Salama, N. R., Shepherd, B. & Falkow, S. Global transposon mutagenesis and essential gene analysis of *Helicobacter pylori*. *J. Bacteriol.* **186**, 7926–7935 (2004).

20. Lamichhane, G., Tyagi, S. & Bishai, W. R. Designer arrays for defined mutant analysis to detect genes essential for survival of *Mycobacterium tuberculosis* in mouse lungs. *Infect. Immun.* **73**, 2533–2540 (2005).

21. Chan, K., Kim, C. C. & Falkow, S. Microarray-based detection of *Salmonella enterica* serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. *Infect. Immun.* **73**, 5438–5449 (2005).

22. Saenz, H. L. & Dehio, C. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. *Curr. Opin. Microbiol.* **8**, 612–619 (2005).

23. Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).

24. Paddison, P. J. *et al.* A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**, 427–431 (2004).
   References 23 and 24 both demonstrate the applicability of STM to RNAi screens, although they use different tagging approaches.

25. Mecsas, J. Use of signature-tagged mutagenesis in pathogenesis studies. *Curr. Opin. Microbiol.* **5**, 33–37 (2002).

26. Autret, N. & Charbit, A. Lessons from signature-tagged mutagenesis on the infectious mechanisms of pathogenic bacteria. *FEMS Microbiol. Rev.* **29**, 703–717 (2005).

27. Coulter, S. N. *et al. Staphylococcus aureus* genetic loci impacting growth and survival in multiple infection environments. *Mol. Microbiol.* **30**, 393–404 (1998).

28. Hava, D. L. & Camilli, A. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* **45**, 1389–1406 (2002).
   This large-scale screen identified some interesting tissue-specific pneumococcal virulence factors.

29. Tsolis, R. M. *et al.* Identification of a putative *Salmonella enterica* serotype Typhimurium host range factor with homology to IpaH and YopM by signature-tagged mutagenesis. *Infect. Immun.* **67**, 6385–6393 (1999).

30. Morgan, E. *et al.* Identification of host-specific colonization factors of *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.* **54**, 994–1010 (2004).

31. Hisert, K. B. *et al.* Identification of *Mycobacterium tuberculosis* counterimmune (cim) mutants in immunodeficient mice by differential screening. *Infect. Immun.* **72**, 5315–5321 (2004).
   In this interesting paper, STM was used in conjunction with immune-deficient mouse strains to identify 'counter-immune' mutants, in which genes with products that protect the pathogen against specific host immune mechanisms are affected.

32. Hisert, K. B. *et al.* A glutamate-alanine-leucine (EAL) domain protein of *Salmonella* controls bacterial survival in mice, antioxidant defence and killing of macrophages: role of cyclic diGMP. *Mol. Microbiol.* **56**, 1234–1245 (2005).

33. Birrell, G. W., Giaever, G., Chu, A. M., Davis, R. W. & Brown, J. M. A genome-wide screen in *Saccharomyces cerevisiae* for genes affecting UV radiation sensitivity. *Proc. Natl Acad. Sci. USA* **98**, 12608–12613 (2001).

34. Birrell, G. W. *et al.* Transcriptional response of *Saccharomyces cerevisiae* to DNA-damaging agents does not identify the genes that protect against these agents. *Proc. Natl Acad. Sci. USA* **99**, 8778–8783 (2002).

35. Game, J. C. *et al.* Use of a genome-wide approach to identify new genes that control resistance of *Saccharomyces cerevisiae* to ionizing radiation. *Radiat. Res.* **160**, 14–24 (2003).

36. Hanway, D. *et al.* Previously uncharacterized genes in the UV- and MMS-induced DNA damage response in yeast. *Proc. Natl Acad. Sci. USA* **99**, 10605–10610 (2002).

37. Deutschbauer, A. M., Williams, R. M., Chu, A. M. & Davis, R. W. Parallel phenotypic analysis of sporulation and postgermination growth in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **99**, 15530–15535 (2002).

38. Deutschbauer, A. M. *et al.* Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925 (2005).

39. Ooi, S. L., Shoemaker, D. D. & Boeke, J. D. A DNA microarray-based genetic screen for nonhomologous end-joining mutants in *Saccharomyces cerevisiae*. *Science* **294**, 2552–2556 (2001).

40. Fleming, J. A. *et al.* Complementary whole-genome technologies reveal the cellular response to proteasome inhibition by PS-341. *Proc. Natl Acad. Sci. USA* **99**, 1461–1466 (2002).

41. Wright, R. *et al.* Parallel analysis of tagged deletion mutants efficiently identifies genes involved in endoplasmic reticulum biogenesis. *Yeast* **20**, 881–892 (2003).

42. Steinmetz, L. M. *et al.* Systematic screen for human disease genes in yeast. *Nature Genet.* **31**, 400–404 (2002).

43. Lum, P. Y. *et al.* Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* **116**, 121–137 (2004).
   Describes the application of a genome-wide barcoding screen in yeast to identify new modes of action for several clinically relevant compounds.

44. Giaever, G. *et al.* Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc. Natl Acad. Sci. USA* **101**, 793–798 (2004).

45. Lee, W. *et al.* Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS Genet.* **1**, 235–246 (2005).

46. Brown, J. A. *et al.* Global analysis of gene function in yeast by quantitative phenotypic profiling. *Mol. Syst. Biol.* **2**, 1–9 (2006).

47. Parsons, A. B. *et al.* Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* **126**, 611–625 (2006).

48. Bender, A. & Pringle, J. R. Use of a screen for synthetic lethal and multicopy suppressee mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **11**, 1295–1305 (1991).

49. Tong, A. H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368 (2001).

50. Tong, A. H. *et al.* Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813 (2004).

51. Ooi, S. L., Shoemaker, D. D. & Boeke, J. D. DNA helicase gene interaction network defined using synthetic lethality analyzed by microarray. *Nature Genet.* **35**, 277–286 (2003).

52. Pan, X. *et al.* A robust toolkit for functional profiling of the yeast genome. *Mol. Cell* **16**, 487–496 (2004).

53. Cormack, B. P. & Falkow, S. Efficient homologous and illegitimate recombination in the opportunistic yeast pathogen *Candida glabrata*. *Genetics* **151**, 979–987 (1999).

54. Las Penas, A. *et al.* Virulence-related surface glycoproteins in the yeast pathogen *Candida glabrata* are encoded in subtelomeric clusters and subject to RAP1- and SIR-dependent transcriptional silencing. *Genes Dev.* **17**, 2245–2258 (2003).

55. Roemer, T. *et al.* Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery. *Mol. Microbiol.* **50**, 167–181 (2003).

56. Brown, J. S. *et al.* Signature-tagged and directed mutagenesis identify PABA synthetase as essential for *Aspergillus fumigatus* pathogenicity. *Mol. Microbiol.* **36**, 1371–1380 (2000).

57. Nelson, R. T., Hua, J., Pryor, B. & Lodge, J. K. Identification of virulence mutants of the fungal pathogen *Cryptococcus neoformans* using signature-tagged mutagenesis. *Genetics* **157**, 935–947 (2001).

58. Moorman, N. J., Lin, C. Y. & Speck, S. H. Identification of candidate gammaherpesvirus 68 genes required for virus replication by signature-tagged transposon mutagenesis. *J. Virol.* **78**, 10282–10290 (2004).

59. Song, M. J. *et al.* Identification of viral genes essential for replication of murine γ-herpesvirus 68 using signature-tagged mutagenesis. *Proc. Natl Acad. Sci. USA* **102**, 3805–3810 (2005).

60. Knoll, L. J., Furie, G. L. & Boothroyd, J. C. Adaptation of signature-tagged mutagenesis for *Toxoplasma gondii*: a negative screening strategy to isolate genes that are essential in restrictive growth conditions. *Mol. Biochem. Parasitol.* **116**, 11–16 (2001).

61. Sontheimer, E. J. Assembly and function of RNA silencing complexes. *Nature Rev. Mol. Cell Biol.* **6**, 127–138 (2005).

62. Westbrook, T. F. *et al.* A genetic screen for candidate tumor suppressors identifies REST. *Cell* **121**, 837–848 (2005).

63. Brummelkamp, T. R. *et al.* An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. *Nature Chem. Biol.* **2**, 202–206 (2006).

64. Silva, J. M. *et al.* Second-generation shRNA libraries covering the mouse and human genomes. *Nature Genet.* **37**, 1281–1288 (2005).

65. Ngo, V. N. *et al.* A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* **441**, 106–110 (2006).
   An excellent example of the fusion of barcoding with RNAi technology, yielding insights into a specific cellular process.

66. Geoffroy, M. C., Floquet, S., Metais, A., Nassif, X. & Pelicic, V. Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. *Genome Res.* **13**, 391–398 (2003).

67. Liberati, N. T. *et al.* An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl Acad. Sci. USA* **103**, 2833–2838 (2006).

68. Nord, A. S. *et al.* The International Gene Trap Consortium Website: a portal to all publicly available gene trap cell lines in mouse. *Nucleic Acids Res.* **34**, D642–D648 (2006).

69. Beutler, B. *et al.* Genetic analysis of host resistance: Toll-like receptor signaling and immunity at large. *Annu. Rev. Immunol.* **24**, 353–389 (2006).

70. Crozat, K. *et al.* Analysis of the MCMV resistome by ENU mutagenesis. *Mamm. Genome* **17**, 398–406 (2006).

71. Grant, A. J. *et al.* Signature-tagged transposon mutagenesis studies demonstrate the dynamic nature of cecal colonization of 2-week-old chickens by *Campylobacter jejuni*. *Appl. Environ. Microbiol.* **71**, 8031–8041 (2005).

72. Sun, D., Melegari, M., Sridhar, S., Rogler, C. E. & Zhu, L. Multi-miRNA hairpin method that improves gene knockdown efficiency and provides linked multi-gene knockdown. *Biotechniques* **41**, 59–63 (2006).

73. Xia, X. G., Zhou, H. & Xu, Z. Multiple shRNAs expressed by an inducible pol II promoter can knock down the expression of multiple target genes. *Biotechniques* **41**, 64–68 (2006).

74. Boucrot, E., Henry, T., Borg, J. P., Gorvel, J. P. & Meresse, S. The intracellular fate of *Salmonella* depends on the recruitment of kinesin. *Science* **308**, 1174–1178 (2005).

75. Kirkpatrick, B. D. *et al.* Evaluation of *Salmonella enterica* serovar Typhi (Ty2 aroC-ssaV-) M01ZH09, with a defined mutation in the *Salmonella* pathogenicity island 2, as a live, oral typhoid vaccine in human volunteers. *Vaccine* **24**, 116–123 (2006).

76. Cox, J. S., Chen, B., McNeil, M. & Jacobs, W. R., Jr. Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. *Nature* **402**, 79–83 (1999).

77. Camacho, L. R., Ensergueix, D., Perez, E., Gicquel, B. & Guilhot, C. Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Mol. Microbiol.* **34**, 257–267 (1999).

78. Sulzenbacher, G. *et al.* LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *EMBO J.* **25**, 1436–1444 (2006).

79. Rousseau, C. *et al.* Production of phthiocerol dimycocerosates protects *Mycobacterium tuberculosis* from the cidal activity of reactive nitrogen intermediates produced by macrophages and modulates the early immune response to infection. *Cell Microbiol.* **6**, 277–287 (2004).

80. West, N. P. *et al.* Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science* **307**, 1313–1317 (2005).

81. Hayes, F. Transposon-based strategies for microbial functional genomics and proteomics. *Annu. Rev. Genet.* **37**, 3–29 (2003).

82. Chiang, S. L. & Mekalanos, J. J. Use of signature-tagged transposon mutagenesis to identify *Vibrio cholerae* genes critical for colonization. *Mol. Microbiol.* **27**, 797–805 (1998).

83. Darwin, A. J. & Miller, V. L. Identification of *Yersinia enterocolitica* genes affecting survival in an animal host using signature-tagged transposon mutagenesis. *Mol. Microbiol.* **32**, 51–62 (1999).

84. Edelstein, P. H., Edelstein, M. A., Higa, F. & Falkow, S. Discovery of virulence genes of *Legionella pneumophila* by using signature tagged mutagenesis in a guinea pig pneumonia model. *Proc. Natl Acad. Sci. USA* **96**, 8190–8195 (1999).

85. Foulongne, V., Bourg, G., Cazevieille, C., Michaux-Charachon, S. & O'Callaghan, D. Identification of *Brucella suis* genes affecting intracellular survival in an *in vitro* human macrophage infection model by signature-tagged transposon mutagenesis. *Infect. Immun.* **68**, 1297–1303 (2000).

86. Martindale, J., Stroud, D., Moxon, E. R. & Tang, C. M. Genetic analysis of *Escherichia coli* K1 gastrointestinal colonization. *Mol. Microbiol.* **37**, 1293–1305 (2000).

87. Sun, Y. H., Bakshi, S., Chalmers, R. & Tang, C. M. Functional genomics of *Neisseria meningitidis* pathogenesis. *Nature Med.* **6**, 1269–1273 (2000).

88. Kavermann, H. *et al.* Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. *J. Exp. Med.* **197**, 813–822 (2003).

89. Claus, H., Frosch, M. & Vogel, U. Identification of a hotspot for transformation of *Neisseria meningitidis* by shuttle mutagenesis using signature-tagged transposons. *Mol. Gen. Genet.* **259**, 363–371 (1998).

90. Lau, G. W. *et al.* A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. *Mol. Microbiol.* **40**, 555–571 (2001).

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
lppX

FURTHER INFORMATION
Boone Laboratory homepage:
http://www.utoronto.ca/boonelab
Centre for Molecular Microbiology and Infection web site:
http://www.imperial.ac.uk/cmmi
The Institute for Genomic Research: http://www.tigr.org
World Health Organisation: http://www.who.int
Access to this links box is available online.