Current Biology, Vol. 12, 863–867, May 14, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0960-9822(02)00837-0

A Simple Method for Genome-Wide Screening for Advantageous Insertions of Mobile DNAs in *Escherichia coli*

Richard J. Edwards, R. Elizabeth Sockett, and John F.Y. Brookfield[1]
Institute of Genetics
University of Nottingham
Queens Medical Centre
Nottingham NG7 2UH
United Kingdom

Summary

Laboratory evolution in *Escherichia coli* has revealed that fitness typically increases in experimental populations [1, 2]. These changes are sometimes associated with changes in insertion sequence positions, some of which may themselves cause advantageous phenotypes [3, 4]. We have a novel and general method for identifying genes in *Escherichia coli*, whose knockout by mobile DNA insertions is beneficial in experimental evolution. Insertion sites in favored clones can be identified by reference to genomic information. We have implemented the method using modified Tn*10* transposons bearing kanamycin and chloramphenicol resistance cassettes. Results are consistent across replicated experiments, demonstrating that the insertions are themselves creating selective advantages, rather than hitch-hiking with favorable base substitutions. The successful clones have subsequently been confirmed to have a fitness advantage relative to the progenitor strain. In experiments in shaking culture, we find that advantageous insertions usually fall in operons required in the pathways creating flagella. The method allows a rapid genome-wide screening for advantageous insertions in arbitrary environmental conditions. It allows investigation of the extent to which transient mutations generating environment-dependent selective advantages may help to explain the persistence of mobile DNAs in primarily clonal organisms, such as *E. coli.*

Results and Discussion

We seek to create a population of cells in which each cell contains a single insertion of a mobile DNA sequence (which will subsequently be stably inherited). However, different cells will have insertions in different and random positions, such that each gene is inactivated in at least some cells in this population. The population is then grown for 60 to 90 generations, and competition between clones will allow high-fitness clones to increase in frequency and be detectable. *E. coli* cells are infected with λ phages bearing mini-Tn*10* transposons, with IS*10* terminal sequences flanking an antibiotic resistance gene, and the IS*10* transposase (with reduced target specificity) supplied by a gene elsewhere on the phage DNA [5]. We have carried out two types of experiments differing in the antibiotic resistance used, with either λNK1316 or λNK1324, in which the mini-Tn*10* element carries resistance to kanamycin or chloramphenicol, respectively. The mini-Tn*10* transposes into the chromosome and is stable thereafter, and cells with insertions are selected by growth on the antibiotic, creating a population of around 10^5 cells with random single insertions. These experimental populations are grown in liquid culture with antibiotic. Populations are subcultured by 100- or 1000-fold dilutions into fresh broth daily, and, from cells sampled from the populations, DNA is extracted, restriction digested, electrophoresed, Southern blotted, and probed with the antibiotic resistance gene's DNA. Figure 1 shows the appearance of discrete fragments in the course of 9 days with daily 100-fold dilutions (or around 60 generations of experimental evolution) using the λNK1316 mini-Tn*10*, conferring kanamycin resistance. Initially, a smear of fragments is seen, representing the randomly distributed insertions in the different clones. However, during the course of the experiment, strongly hybridizing fragments accumulate at discrete points. These are the fragments bearing Tn*10* insertions in the high-fitness clones. The results were consistent across replicates (Figure 2). (Experiments in which the daily dilution was 1000-fold yielded the same fragments but more rapidly.)

Different fragments represent different clones, and, potentially, all mobile DNA insertions that create significant fitness increases can be detected, rather than merely the one creating the greatest fitness increase (which would have been detected had the most advantageous clone been allowed to spread to fixation).

Single clones sampled from the final populations had their DNAs extracted, which was digested, electrophoresed, Southern blotted, and probed. DNAs were also digested, ligated into pBR322, and PCR amplified using primers from the plasmid and from the IS10R termini. The sequence obtained was used to screen the *E. coli* K-12 genome sequence [6] and the gene knocked out in the high-fitness clone and the position of the insertion in the gene were identified. The size of the restriction fragment bearing the Tn*10* insertion was used as a guide to whether the clone corresponded to one of the high-frequency clones observed in the final generation of selection.

What is the mechanism for the generation of high fitness by insertions of the modified Tn*10*? One possibility is that it might be related to the selection for antibiotic resistance imposed throughout the experiment. Thus, we have repeated the experiment with λNK1324, which carries a mini-Tn*10* encoding chloramphenicol resistance (see Experimental Procedures). Autoradiograms documenting the evolution of populations selected on chloramphenicol (data not shown) again revealed specific restriction fragments appearing in the course of selection, representing the restriction fragments, insertions into which create high-fitness clones.

We have identified a total of 40 genes as having insertion sequences at the end of the experiment: 24 of which

[1]Correspondence: john.brookfield@nottingham.ac.uk

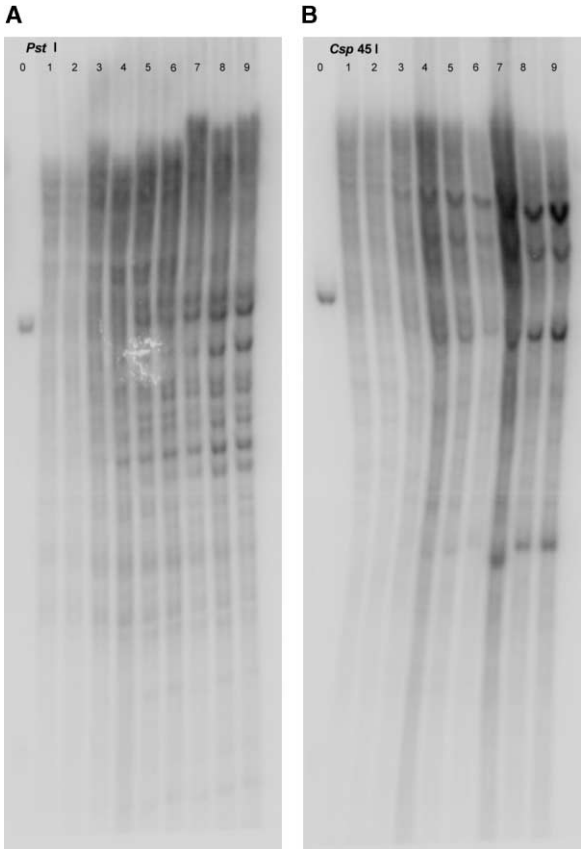

Figure 1. Autoradiograms of Experimental Populations Probed with Labeled *kan*^{R} DNA

Lanes 0 to 9 represent populations of mixtures of transformed *E. coli* strain W3110 genomic DNAs digested with PstI (A) and Csp45I (B), which do not cut within the mini-Tn*10* transposon, after 0–9 days, respectively, of growth with shaking in 50 ml of luria broth with 30 μg/ml kanamycin. Dilution is 100-fold daily. The fragment at generation zero represents the Tn*10* still on the phage vector in abortive lysogens, circularized phage that confer resistance but cannot reproduce. The bands appearing from 4 to 9 days of growth represent mini-Tn*10* insertions in high-fitness clones.

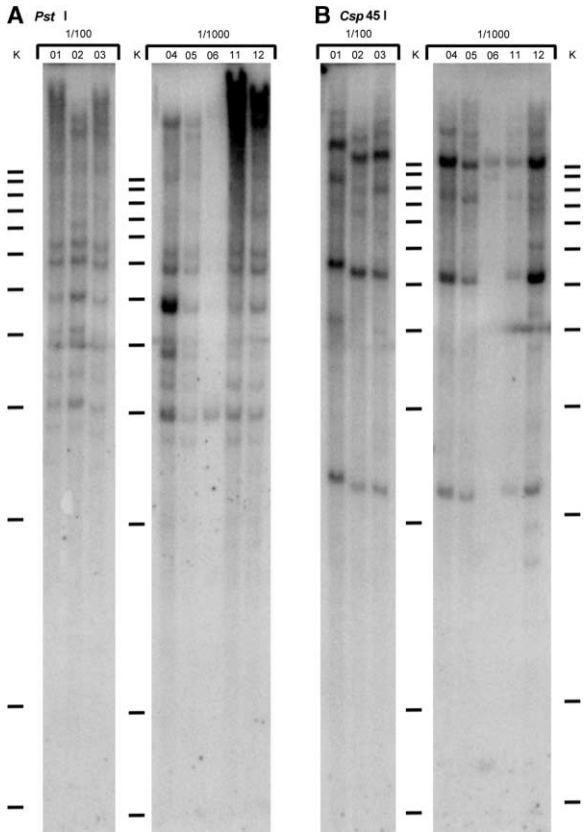

Figure 2. Autoradiogram Showing Repeatability of the Final Outcome with Experiments Using Kanamycin Selection with Dilutions of 100-Fold and 1000-Fold

(A) Shows final generation populations digested with PstI and (B) the same populations digested with Csp45I. Each shows three populations diluted 100-fold daily and five populations diluted 1000-fold daily. One population only (in lane 6) shows a reduced number of bands. This population experienced an unexpected population bottleneck during growth.

were identified in the experiments with kanamycin selection and 22 identified in experiments with chloramphenicol selection. (Six genes were identified in both sets of experiments.) (These are shown in Table S1 in the Supplementary Material available with this article online.) Some of these genes have been identified as corresponding to the strong restriction fragments identified on the Southern blots. There are ten such genes in the kanamycin experiments, and eleven in the chloramphenicol experiments, and all six of the genes detected in both experiments are included. Among the genes detected as having insertions, 17, including all six detected in both the kanamycin and chloramphenicol experiments, are in operons required for the production of flagella. Electron micrographs of the progenitor cell W3110 showed it to possess flagella, although these appear to be paralyzed, the cause of which is unknown. A clone from the final generation with a mutation in the flagellar export-switch gene *fliM* now lacked flagella, as did clones with mutations in the *flgE* hook and *fliR* flagellar export genes. Of interest is a clone which has an insertion into the previously poorly characterized gene *gabP*. There are no flagella on this cell (although apparently pilus-like structures are seen around the poles).

Previous reports have indicated that *E. coli* cells lacking flagella have an advantage in liquid culture [7]. This suggests that our insertion mutations are truly advantageous, rather than representing rare sites in which insertions are neutral. Furthermore, it would suggest that this advantage is not the result of the paralysis of the flagella in the starting bacteria. To confirm that the insertions truly are advantageous relative to the starting strain, we competed clones derived from the end of the experiment against the progenitor clone, W3110, in the same environmental conditions as were used in the initial experiments, save for the absence of the antibiotic in this test (since this would kill the progenitor). Earlier results [8] indicate that in the absence of antibiotic, the insertion elements used are approximately selectively neutral. The results shown in Table 1 confirm that flagella gene insertions are advantageous, not neutral, and that the advantage is not related to generation of antibiotic resis-

Table 1. Estimated Fitnesses of Clones Sampled from the End of Experiment, Measured in Competition with Progenitor Genotype W3110

Mutant[a]	Percent Resistant Prior to Selection[b]	Percentage Resistant after Selection[b]	Selective Advantage of Mutant[c], *s*
flgE	58.3 ± 3.2	77.7 ± 14.1	(0.009−) **0.096** (−02.48)
fliE/F	51.8 ± 2.4	86.9 ± 5.9	(0.138−) **0.200** (−0.295)
fliM (1)	49.8 ± 3.6	76.3 ± 3.0	(0.092−) **0.126** (−0.163)
fliM (2)	60.8 ± 3.4	74.2 ± 4.0	(0.028−) **0.064** (−0.103)
fliR (1)	51.9 ± 4.2	84.1 ± 13.1	(0.067−) **0.173** (−0.442)
fliR (2)	52.8 ± 3.5	91.6 ± 19.3	(0.073−) **0.257** (−∞)
b1478	46.8 ± 2.8	46.2 ± 3.4	(−0.027−) **0.002** (−0.023)
gabP	47.6 ± 3.2	68.1 ± 3.2	(0.058−) **0.089** (−0.124)
yojN (1)	30.3 ± 2.1	57.4 ± 4.9	(0.088−) **0.120** (−0.155)
yojN (2)	53.3 ± 2.3	85.2 ± 5.9	(0.119−) **0.176** (−0.257)
b2461	46.7 ± 2.5	39.5 ± 4.5	(−0.057−) **−0.029** (− −0.001)
b2461	50.7 ± 1.9	45.1 ± 2.4	(−0.039−) **−0.022** (− −0.005)
yjhS	51.0 ± 3.0	46.0 ± 3.2	(−0.044−) **−0.020** (− −0.005)

[a] Gene disrupted in the insertion mutation tested.
[b] Percentage of resistant cells before and after competition with wild-type (± SE).
[c] Estimated selection coefficient in favor of the mutant (in bold), calculated as in Experimental Procedures. The range of values is calculated by adding or subtracting one standard error from the percentages resistant at the start and end of the experiment.

tance. The fitness of the clones can be calculated from this experiment as shown in the Experimental Procedures, and estimated fitness values are also shown in Table 1. Of the clones tested in this way, six have insertions in four different known flagellar genes, and seven other tested clones have insertions in five other genes. All of the flagellar gene insertions have fitness advantages (with estimates of *s* ranging from 6.4% to 25.7%) relative to the progenitor clone W3110. Of the remainder, an insertion in *b1478*, which retains flagella but which is a candidate for being a high-frequency insertion site in the chloramphenicol experiment only, did not create a fitness advantage relative to the starting strain. Possibly, insertions here are advantageous only under chloramphenicol selection. Insertions in genes *gabP* (from the kanamycin experiment) and *yohN* (from the chloramphenicol experiment) were advantageous. *yojN* has not been tested for the presence of flagella.

The tested *b2461* and *yjhS* clones were sampled randomly from the final generation of a chloramphenicol

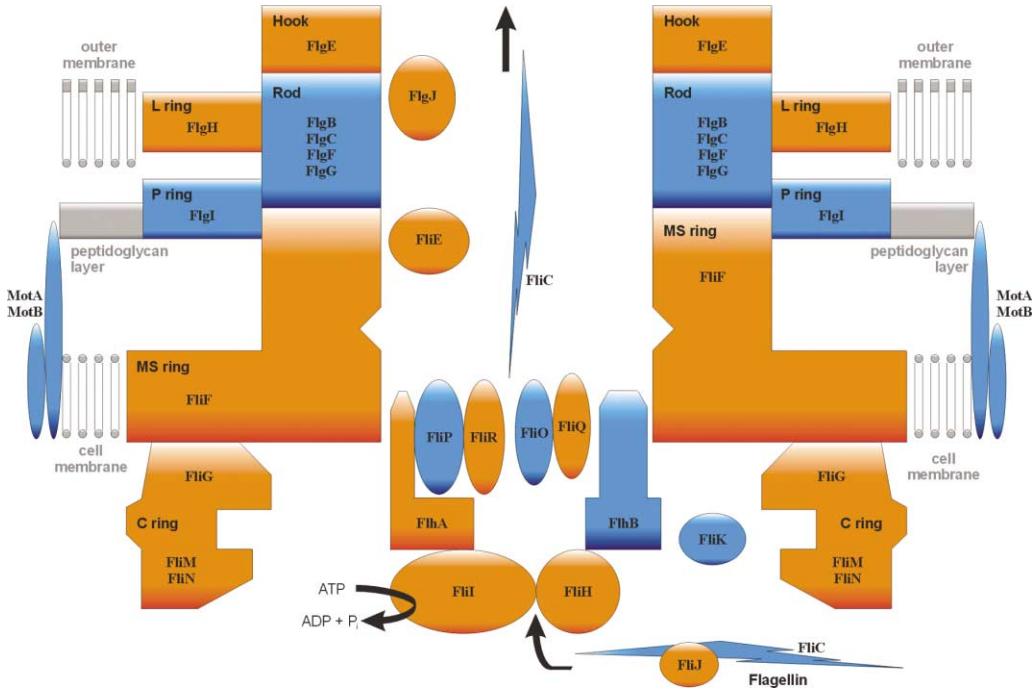

Figure 3. Locations and Roles of the Proteins Disrupted in High-Frequency Clones in Flagella Production

The blocks shown in orange represent groups of proteins including at least one of which has been disrupted in one of the high-frequency clones. Those shown in blue do not have insertions among the high-frequency clones identified in our restricted study of advantageous mutation.

experiment, and insertions in these genes are not responsible for any of the high-frequency fragments in the autoradiogram. It is possible that these have persisted as weakly advantageous clones throughout the experiment, although, no doubt, diminishing in frequency in the later stages in competition with the favored clones. Not all of the clones creating strongly hybridizing fragments in Figures 1, 2, and 3 have been tested for having improved fitnesses. However, given that some of the clones appearing have elevated fitnesses, it can be inferred that the remaining clones creating strong fragments in these autoradiograms cannot be merely neutral, since they would then have been out-competed by the high-fitness clones. Simulations reveal that selective advantages of 4% or more should be sufficient to create visible bands in the course of 60 to 90 generations of population growth. Thus, it is unsurprising that our advantageous clones, with *s* values typically two to six times this, were able to spread to high frequencies in this time.

The cells studied in the experiments will be subject to base substitution mutations in addition to the transposon insertions. We believe that the cells are not significantly adapting by the spread of base substitution mutations during the course of the experiment. For this to happen, the mutations generated would have to have even higher fitnesses than the transposon insertions that we are creating. If such advantageous base substitutions were occurring, they would arise in cells with particular transposon insertions and, as they spread, would also spread the particular transposon insertion that existed in the favored cell. There would thus be transposon insertions that would spread to high frequencies in the experiment, as a result of hitch-hiking. These insertions would be expected to be in different genes in different replicates of the experiment. However, we find almost complete repeatability of the transposon insertion pattern of the population of cells created during the experiment (Figure 2). This implies that there is little evolution going on by base substitution mutations. The absence of flagella in some advantageous insertion mutations implies that one contributor to the increased fitness could be the avoidance of the production of the flagellin protein encoded by *fliC*. Each flagellum contains around 20,000 subunits of this protein, indicating that flagella loss could be associated with substantial saving of energy and amino acids. The roles of the flagellar proteins whose genes are disrupted in these experiments are shown on Figure 3.

This method could potentially be applied to *E. coli* evolution in any environment. Antibiotic selection is required at the outset to ensure that all of the cells have a Tn*10* insertion but afterwards could be relaxed. With the more than 4000 genes in the *E. coli* genome having evolved in complex environments, it seems highly probable that in most simple or novel environments there will be no requirement for many of the genes of the *E. coli* genome, and the inactivation of some of these will indeed benefit the cell.

The discovery of advantageous sites is also of great general interest, since the maintenance of mobile DNAs in clonal populations is hard to explain unless insertions are sometimes advantageous [9]. While *E. coli* in the wild have opportunities for genetic exchange while our laboratory populations do not, advantageous mutations may nevertheless play an important role in some environments. We have modeled (R.J.E. and J.F.Y.B, submitted) the influence of selection for transposon insertion mutations in a fluctuating environment and shown that such selection can maintain a family of mobile DNA sequences over a broad parameter space.

Experimental Procedures

Insertions of mini-Tn*10* constructs were made using λNK1316 (kanamycin resistance gene from Tn903) and λNK1324 (chloramphenicol resistance from pACYC184) [5]. Infection introduces these into the cell on modified phage DNAs bearing an IS*10*-derived transposase with altered target specificity driven by a Ptac promoter responding to 1 mM IPTG. *E. coli* W3110 [F^-,λ^-, IN(*rrnD-rrnE*)1] were grown overnight (in 50 ml tryptone broth with 0.2% w/v maltose and 10 mM Mg_2SO_4) at 37°C to 10^9–10^{10} cells/ml, then concentrated into 5 ml luria broth (per l:10 g bactotryptone, 5 g bactoyeast extract, 10 g NaCl). Phage stock (500 μl) (10^9–10^{10} pfu/ml) was added and the culture incubated for 15 min at room temperature, then a further 15 min at 37°C. Aliquots (5 × 1 ml) were added to 5 ml luria broth with 50 mM sodium citrate, centrifuged, and resuspended in 5 ml luria broth with 50 mM sodium citrate. IPTG was added to 1 mM and antibiotic added (30 μg/ml for kanamycin, 25 μg/ml for chloramphenicol). Cultures were incubated overnight at 42°C and then combined into one culture (labeled "Generation 0"). Each day, lines were diluted either 100- or 1000-fold into 50 ml luria broth plus antibiotic and incubated overnight at 37°C. This was repeated nine times, and samples were taken each day, giving ten samples, "Generations 0–9." Bacterial DNA was digested with restriction enzymes, separated electrophoretically, and analyzed by Southern hybridization following standard procedures [10].

Gel extracted restriction fragments from individual generation nine clones were ligated to vector DNAs (pBR322 or pBluescript SK+). PCR was performed using a primer binding to vector sequence and one from the antibiotic resistance gene, and PCR products were sequenced to identify the mini-Tn*10* insertion site.

Fitness assays were performed without antibiotic using 50 μl of a 50:50 mixture of an overnight culture of W3110 and that of the strain under test. The mixture was kept in luria broth for 24 hr. Prior to and subsequent to this growth phase, 50 μl samples of 10^{-6} dilutions were spread on 12 plates with and without the relevant antibiotic (giving ~100–200 colonies per plate). The ratios of antibiotic-resistant and antibiotic-sensitive colonies before and after the growth phase were used to calculate fitness as follows.

If *x* represents the proportion of cells that are resistant at the end of competitive growth, and *y* the proportion at the start, then the relative fitness of the clone under test over the ten generations of competition, ω, is given by $x(1 - y)/(y(1 - x))$. The selective advantage, *s*, is calculated by $1 + s$ being the tenth root of ω. The twelve replicate plates in each experiment allow the calculation of the standard errors of the ratios shown in Table 1.

Electron Microscopy

Cells from mid-log phase populations of *E. coli* were applied to carbon-coated formvar grids (Agar Scientific) and washed briefly with 10 mM Tris 0.1 mM EDTA (pH 8). They were stained for 30 s with 1% uranyl acetate (pH 5.5) and examined using a JEOL JEM-100S electron microscope.

Supplementary Material

Supplementary Material including a figure (Figure S1: Southern Hybridizations Showing Results of the Chloramphenicol Selection Experiment, with Populations Sampled after 0, 1, 3, 5, 7, and 9 Days of Competitive Growth) and two tables (Table S1: Genes Identified as Possessing Insertions at the End of the Experiment Using Kanamycin Selection; Table S2: Genes Identified as Possessing Insertions at the End of the Experiment Using Chloramphenicol Selection) can be found online at http://images.cellpress.com/supmat/supmatin.htm.

Acknowledgments

We thank Maggie Smith and John Armour for useful discussion and we thank the University of Nottingham for financial support for R.J.E.

Received: January 25, 2002
Revised: March 27, 2002
Accepted: March 27, 2002
Published: May 14, 2002

References

1. Lenski, R.E., Rose, M.R., Simpson, S.C., and Tadler, S.C. (1991). Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations. Am. Nat. *138*, 1315–1341.
2. Travisano, M., Mongold, J.A., Bennett, A.F., and Lenski, R.E. (1995). Experimental tests of the roles of adaptation, chance and history in evolution. Science *267*, 87–90.
3. Papadoupolos, D., Schneifder, D., Meier-Eiss, J., Arber, W., Lenski, R.E., and Blot, M. (1999). Genomic evolution during a 10,000 generation experiment with bacteria. Proc. Natl. Acad. Sci. USA *96*, 3807–3812.
4. Schneider, D., Duperchy, E., Coursange, E., Lenski, R.E., and Blot, M. (2000). Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. Genetics *156*, 477–488.
5. Kleckner, N., Bender, J., and Gottesman, S. (1991). Uses of transposons with emphasis on *Tn10*. Methods Enzymol. *204*, 139–180.
6. Blattner, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. Science *277*, 1453–1462.
7. McNab, R.M. (1987). Flagella in Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology, F.C. Neidhardt, Ed. (Washington, DC: American Society for Microbiology), pp. 70–83.
8. Elena, S.F., Ekunwe, L., Hajela, N., Oden, S.A., and Lenski, R.E. (1998). Punctuated evolution caused by selection of rare beneficial mutations. Genetica *102*, 349–358.
9. Hickey, D.A. (1982). Selfish DNA: a sexually-transmitted nuclear parasite. Genetics *101*, 519–531.
10. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning. A Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Press).