

Identifying Essential *Streptococcus sanguinis* Genes Using Genome-Wide Deletion Mutation

Lei Chen, Xiuchun Ge, and Ping Xu

Abstract

Essential genes in pathogens are important for the development of antibacterial drugs. In this report, we described a protocol to identify essential genes in the *Streptococcus sanguinis* SK36 strain using genome-wide deletion mutation. A fusion PCR-based method is used to construct gene deletion fragments, which contain kanamycin resistance cassettes with two flanking arms of DNA upstream and downstream of the target gene. The linear fused PCR amplicons were transformed into *S. sanguinis* SK36 cells. No kanamycin-resistant transformants suggested the gene essentiality because the deletion of the essential gene renders a lethal phenotype of the transformants. The putative essential genes were further confirmed by independent transformations up to five attempts. The false nonessential genes were also identified by removing double-band mutants.

Key words Essential genes, *Streptococcus sanguinis*, Fusion PCR, Deletion mutation, Genome-wide

1 Introduction

Essential genes are defined as those that are indispensable for cellular life [1, 2]. Identification of essential bacterial genes is important for the development of antibacterial drugs, for designing new bacterial strains for synthetic biology, and for understanding the origins of life [2]. However, systematic identification of essential genes remains challenging because of various limitations. *Streptococcus sanguinis* is a member of human indigenous oral microflora that forms dental plaques, and is one of the most recognized agents of infective endocarditis [3–5]. The transformation frequency of *S. sanguinis* is high [6] and its relatively small genome has been completely sequenced [7]. Hence, *S. sanguinis* is an ideal candidate for genome-wide identification of essential genes. We performed genome-wide essential gene identification and identified a clearly defined set of essential genes in *S. sanguinis* [2, 8]. In this study, a fusion PCR-based method was used to construct gene deletion fragments that contain a kanamycin resistance cassette

with flanking arms of the upstream and downstream DNA of the target gene. Subsequently, the linear fused PCR amplicon was used for direct transformation into *S. sanguinis* SK36. Genes were identified as essential by either absence of the kanamycin-resistant transformants or presence of double-band transformants after five attempts. We finally identified 218 essential genes from a total 2,266 ORFs. Among these, 60 yielded no transformants and 158 were double-band transformants [2, 8]. These essential genes resulted in a simplified model for gene essentiality in *S. sanguinis*, which will greatly facilitate the prediction and validation of essential genes in other pathogenic species.

2 Materials

Prepare all solutions using ultrapure water and analytical grade reagents, and store all prepared reagents at room temperature unless indicated otherwise. Follow all waste disposal regulations when disposing waste materials.

1. PCR components: High Fidelity PCR buffer, 10 mM dNTP mixture, 50 mM MgSO₄, and Platinum Taq High Fidelity DNA polymerase; Mix the required components and store in PCR tubes placed on ice.
2. 96-well PCR plate.
3. 50× TAE buffer: 2.0 M Tris–acetic acid and 50 mM EDTA (pH 8.0).
4. Ethidium bromide and 10 mg/ml stock solution in water; working concentration, 0.5 µg/ml.
5. 48-well or 96-well high-throughput agarose gel electrophoresis system.
6. Todd Hewitt (TH) broth: adjust pH to 7.6 using 10 N NaOH, heat to boiling point, then cool to room temperature and sterilize using 0.22-µm polystyrene filters. Before use, add 300 µl of heat-inactivated horse sera (HS) to a final concentration to 2.5 % in 11.7 ml TH broth in 15-ml conical tubes to obtain TH + HS medium.
7. 35 ng/µl *S. sanguinis* SK36 competence stimulating peptide [9].
8. Brain heart infusion (BHI) agar/broth.
9. 500 mg/ml stock kanamycin solution.
10. Mart anoxomat system [10, 11].

3 Methods

3.1 Primer Design and Construction of Gene Deletion Fragments

1. Primers are designed on the basis of the *S. sanguinis* SK36 genome sequence. Three sets of primers: F1/R1, F2/R2, and F3/R3 are designed for amplification of approximately 1 kb upstream sequence of the target gene, the *aphA-3* encoding kanamycin resistance (Km^r) protein, and approximately 1 kb downstream sequence of the target gene (*see Note 1*), respectively (Fig. 1). Primers R1 and F3 are designed to contain 25 bp adaptor sequences at their 5' end that are complementary to the *aphA-3*. The primers are designed such that their melting temperatures would be as close as possible to 60 °C, facilitating a uniform annealing temperature for all PCR reactions in 96-well plates (*see Note 2*).
2. Primers F1, R1, F3, and R3 are ordered in 96-well plates based on gene order. Each plate contains one type of primers in the same gene order.
3. Dilute the primers in 96-well plates to a final concentration of 10 μ M for PCR amplification using a multichannel pipette to obtain the working primer plate (*see Note 3*).
4. PCR amplification: For the *aphA-3* (*see Note 4*), perform large scale PCR amplification only once to create single-gene deletion constructs. Digest plasmids containing Km^r cassettes using

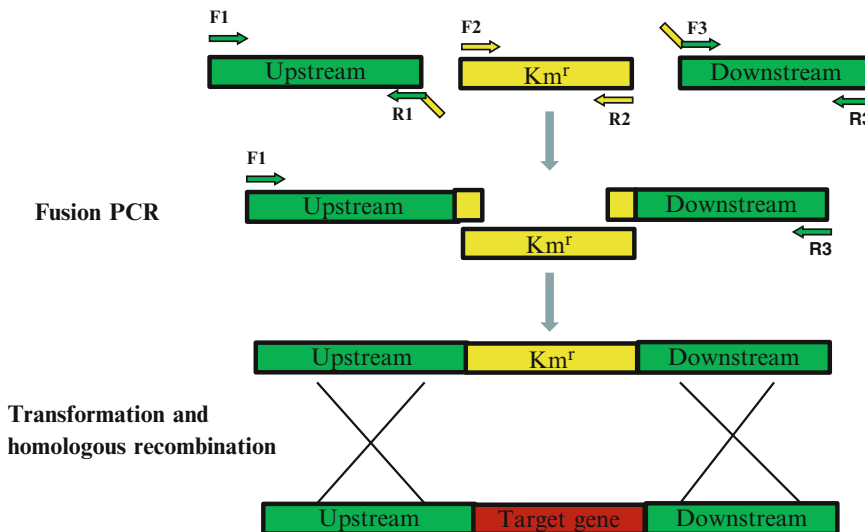


Fig. 1 Schematic construction of the gene deletion fragment by fusion PCR. PCRs are performed using F1/R1, F2/R2, and F3/R3 primers. The three amplicons are purified and mixed in equal amounts as template and a fusion PCR amplicon is obtained using F1/R3, which is used for the following transformation and homologous recombination

EcoRI to serve as a PCR template (*see* **Note 5**). Purify digested plasmid using a QIAquick PCR purification kit and dilute the linear plasmid DNA to a final concentration of 10 ng/ μ l. Prepare 25 μ l mixtures for Km^r cassette amplicons by adding 16.4 μ l ddH₂O, 2.5 μ l 10 $\times$ High Fidelity PCR buffer, 2 μ l of 10 mM dNTP mixture, 1 μ l of 50 mM Mg SO₄, 1 μ l of 10 μ M F2, 1 μ l of 10 μ M R2, 1 μ l of linear plasmid DNA, and 0.1 μ l of Platinum Taq High Fidelity DNA polymerase. Perform PCR amplification at 94 °C for 1 min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 1 min. Examine PCR amplicon using electrophoresis on 1 % agarose gels. Pool Km^r cassette amplicons from ten individual PCR reactions after purification using a QIAquick PCR purification kit. Adjust amplicon concentrations to 10 ng/ μ l.

5. To amplify approximately 1 kb upstream or downstream regions of target *S. sanguinis* genes, prepare PCR mixtures on ice in 15-ml conical tubes containing 1,640 μ l ddH₂O, 250 μ l 10 $\times$ High Fidelity PCR buffer, 200 μ l of 10 mM dNTP mixture, 100 μ l of 50 mM MgSO₄, 100 μ l of 10 ng/ μ l *S. sanguinis* SK36 genomic DNA, and 10 μ l Platinum Taq High Fidelity DNA polymerase, and transfer 23 μ l of the mixture into each well of a 96-well PCR plate using a multichannel pipette. Transfer 1 μ l of corresponding F1 and R1 (or F3 and R3) primers from 10 μ M working primer plates to the PCR plate using a multichannel pipette. Seal the PCR plate and perform amplification at 94 °C for 1 min, followed by 30 cycles of 94 °C for 30 s, 54 °C for 30 s, and 68 °C for 1.5 min.
6. DNA agarose electrophoresis: Prepare 1 % agarose gels containing ethidium bromide with 48 wells to allow for multichannel pipette loading. Mix 4 μ l of PCR amplicon with 1 μ l of 5 $\times$ DNA loading buffer in 96-well plates and load the samples onto agarose gels using a 10- μ l multichannel pipette. Run electrophoresis in 1 $\times$ TAE buffer at 135 V for 30 min. Check bands on gels according to the UVP documentation and analysis system (*see* **Note 6**).
7. PCR amplicon purification: Purify PCR amplicons using a PureLink 96 PCR purification kit, and centrifugation according to the manufacturer's instructions. To elute DNA, add 30 μ l of sterile ddH₂O to the wells of binding plates. Randomly select several purified amplicons on the plate and check DNA concentrations using a NanoDrop spectrophotometer. Adjust the concentration of amplicons on the plate to about 10 ng/ μ l.
8. Fusion PCR: To obtain the final linear recombinant PCR amplicon, combine 1 μ l aliquots of the three PCR fragments from above (in nearly equal-molar amounts) in single PCR plate wells as the template. Add 14.4 μ l ddH₂O, 2.5 μ l 10 $\times$ High Fidelity PCR Buffer, 2 μ l of 10 mM dNTP mixture,

1 μ l of 50 mM Mg SO₄, 1 μ l of 10 μ M F1, 1 μ l of 10 μ M R3, and 0.1 μ l Platinum Taq High Fidelity DNA polymerase. Perform PCR reactions at 94 °C for 2 min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 3.5 min, and final extension at 68 °C for 4 min. Check the PCR amplicons on agarose gels. Purify and quantify the PCR amplicons as described above. Freeze the recombined PCR amplicons at -20 °C.

3.2 Cell Transformation

1. Prepare fresh TH+HS medium and aliquot into 2- and 10-ml tubes.
2. Inoculate 5 μ l of stock *S. sanguinis* SK36 (stored at -80 °C) in 2-ml TH+HS medium, cap tightly and culture overnight at 37 °C, and pre-incubate 10-ml TH+HS tubes concomitantly.
3. After incubation overnight, transfer 50 μ l aliquots of cultures into 10-ml TH+HS tubes and incubate at 37 °C for 3 h to obtain OD660 of 0.07–0.08, and use immediately for transformations.
4. Cell transformation: Add 2 μ l of 70 ng *S. sanguinis* SK36 competence stimulating peptide and 2 μ l of linear recombinant PCR amplicon (approximately 50 ng) to Eppendorf tubes on 96-well blocks and pre-warm at 37 °C. Transfer 330 μ l of SK36 cultures incubated for 3 h in each tube. Incubate at 37 °C for 1 h. Replace DNA with sterile ddH₂O as a control (*see Note 7*).
5. Place the block on ice and spread 100 μ l of each transformation on BHI agar plates (*see Note 8*) containing 500 μ g/ml kanamycin. Incubate the plates at 37 °C for 2 days under microaerobic conditions.

3.3 Identification of Essential Genes

1. Check transformants on kanamycin plates. If no transformants are present, the target gene is identified as an essential gene candidate.
2. If transformants are present on the kanamycin plates, record the approximate number of total colonies on each plate (*see Note 9*). Randomly select two individual colonies and check if they are correct mutants using the following protocols.
3. Randomly select two individual colonies, inoculate into 5 ml of BHI containing 500 μ g/ml kanamycin, and incubate the inoculum microaerobically overnight at 37 °C (*see Note 10*).
4. To determine whether the mutants contain the expected gene replacements, perform colony PCR for each mutant using F1 and R3 primers in 96-well PCR plates. Use 1 μ l aliquots from overnight cultures of individual colonies as the DNA template in 25 μ l PCR reactions. Perform PCR at 94 °C for 5 min, followed by 32 cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 3.5 min, and final extension at 68 °C for 4 min.

5. To accurately identify double-band mutants or contaminant PCR amplicons, examine the PCR amplicons by electrophoresis for 4 h using >12 cm long 1 % agarose gels containing ethidium bromide. Under these electrophoresis conditions, amplicons with ≥ 100 bp differences can be clearly identified. When bands from the Km^r cassette and the wild-type gene are anticipated to differ by <100 bp, an internal T1 primer is used to determine whether the wild-type gene fragment can be detected by PCR (*see Note 11*).
6. To confirm correctness of the deletion, purify amplicons using a PureLink 96 PCR purification kit, and then sequence using the P1 primer for the Km^r cassette.
7. Target genes are identified as essential after five attempts if (1) no transformants are obtained at **step 1**, or (2) if both the selected colonies are confirmed as double-band mutant or a wild-type gene fragment can be detected by PCR at **step 5**, or (3) the sequencing results reveal that the deletion is incorrect as expected at **step 6** (Fig. 2; *see Note 12*).

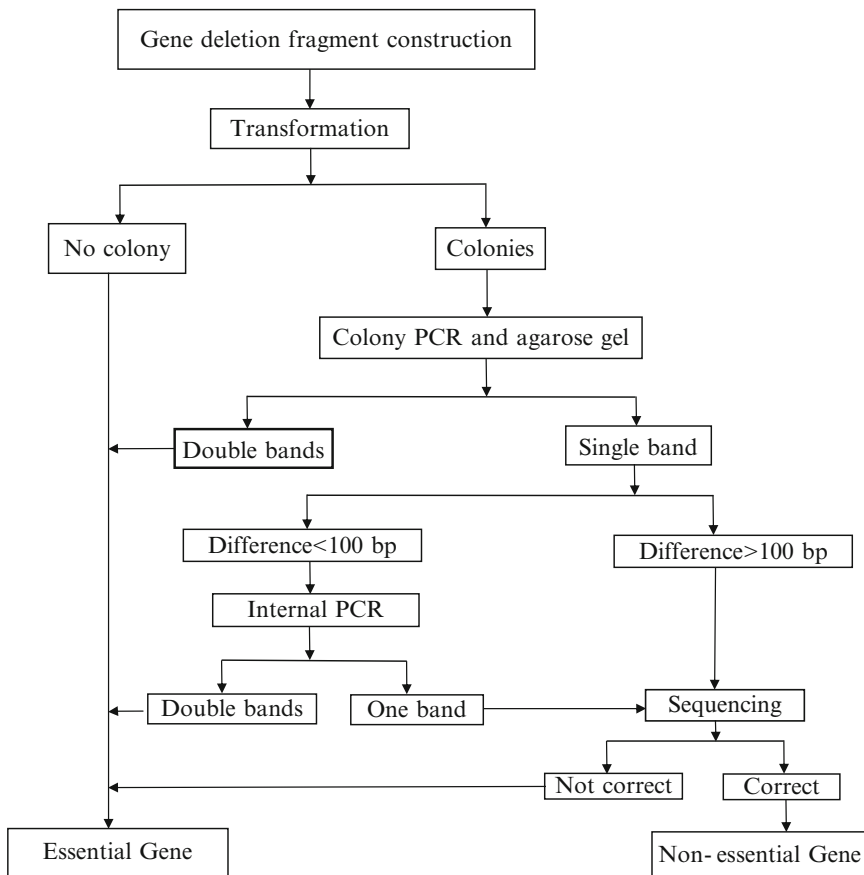


Fig. 2 Flowchart of essential genes identification

4 Notes

1. Lengths of flanking sequences upstream and downstream of the target gene may vary. In this study, flanking sequences were limited to approximately 1 kb based on the following factors: (a) the total length of the PCR-amplifiable amplicon (approximately 3 kb: 1 kb for the *aphA-3* and 2 kb of flanking sequences) and (b) on the feasibility of sequencing flanking regions from both ends (approximately 1 kb) by Sanger's method.
2. The specificity of PCR primers is critical. We found that high primer melting temperatures ($>58^{\circ}\text{C}$), longer primer sizes (>21 bp) and unique primer binding sites in the genome of *S. sanguinis* are important. In addition, it is crucial that only a single gene-specific product is obtained from the final PCR amplification [8].
3. When diluting the primers in 96-well plates using a multichannel pipette, it is important to avoid cross-contamination between neighboring wells. Thus, discard the 96-well plate sealing film after each opening to avoid possible cross-contamination.
4. For studies that aim to identify essential genes, it is preferable to include the native promoter of the *aphA-3* in the deletion construct to address instances of poor expression of the *ahpA-3*. For studies of functional mutant constructs, a promoterless Km^r cassette is preferred to eliminate possible polar transcriptional effects on neighboring genes.
5. A thorough digestion using *EcoRI* is important to avoid possible false positive transformants in the following cell transformation steps.
6. If there are no PCR amplicon or the amplicon is not a single band in one or more wells of the 96-well plate, each PCR reaction must be optimized individually.
7. Negative controls will show transformants if wild type *S. sanguinis* SK36 samples are contaminated with kanamycin resistant mutants. In this case, use backup *S. sanguinis* SK36 stocks and repeat the cell transformation steps.
8. Essential genes are defined as those with lethal consequences under certain environmental conditions [2]; therefore, essential genes are conditional. Differences in screening conditions may cause differences in essential gene identification. As we found for *S. sanguinis*, the use of minimal medium results in identification of additional essential genes compared with use of BHI medium.
9. Genes can be falsely identified as essential because of low transformation frequency. Hence, high transformation frequencies are critical for identification of essential genes. After optimization of the *S. sanguinis* SK36 transformation efficiency, up to

2×10^6 mutant colonies could be obtained from 10^7 bacterial cells with nonessential gene transformations [6, 8], indicating that *S. sanguinis* SK36 is an ideal system for identification of essential genes. We found that essential gene deletions are almost always accompanied by low colony numbers (or no colony at all) on kanamycin plate after cell transformation, although low numbers of colonies do not necessarily indicate that a gene is essential. In this case, inclusion of a control that indicates high transformation frequency in every experiment is preferred. Therefore, nonessential gene deletion constructs that are known to show over 1,000 colonies on the kanamycin plates are used (we used *ssa_0169* [11]).

10. Some transformants demonstrate tiny colonies on kanamycin plates and grow slowly in liquid BHI medium. In this case, incubate the inoculum for another day under the same conditions. In most cases, poor viability indicates that the deleted gene is important but not essential.
11. Use wild-type SK36 as a positive control strain for this PCR. If a band with the expected size is amplified from the mutant using the internal primer, the mutant is identified as a double-band because the presence of kanamycin cassette in the mutant is confirmed by kanamycin selection and subsequent DNA sequencing using the P1 primer.
12. In this study, two types of essential genes were identified in *S. sanguinis*. Deletion of the first type yielded no transformants. Thus, new amplicons for these genes were re-amplified from the beginning and were re-transformed in a second cycle. Subsequently, 60 genes in *S. sanguinis* SK36 were not successfully deleted after five independent attempts, and were classified as essential. The second category of essential genes was identified by mutant colonies producing double-bands in PCR amplifications using F1 and R3 flanking primers. In general, one of the bands corresponded to the size expected for the replacement mutant, whereas the other matched the wild-type gene. We interpreted this as an indication that these genes were also essential because selection for kanamycin resistance resulted in duplication of the target gene [12]. After five independent attempts, we identified a total of 158 double-band essential genes.

Acknowledgment

The authors thank Xiaojing Wang, Yuetan Dou, Jerry Z. Xu, Jenishkumar R Patel, Victoria Stone, My Trinh, Karra Evans, Todd Kitten, Danail Bonchev, and Gregory A. Buck for their contributions in this research. The research is supported by grants R01DE018138 (PX) and R01DE023078 (PX) from the National Institutes of Health.

References

1. Juhas M, Eberl L, Church GM (2012) Essential genes as antimicrobial targets and cornerstones of synthetic biology. *Trends Biotechnol* 30:601–607
2. Xu P, Ge X, Chen L, Wang X, Dou Y et al (2011) Genome-wide essential gene identification in *Streptococcus sanguinis*. *Sci Rep* 1:125
3. Chen L, Ge X, Wang X, Patel JR, Xu P (2012) SpxA1 involved in hydrogen peroxide production, stress tolerance and endocarditis virulence in *Streptococcus sanguinis*. *PLoS One* 7(6):e40034
4. Nyvad B, Kilian M (1987) Microbiology of the early colonization of human enamel and root surfaces in vivo. *Scand J Dent Res* 95:369–380
5. Mylonakis E, Calderwood SB (2001) Infective endocarditis in adults. *N Engl J Med* 345:1318–1330
6. Paik S, Senty L, Das S, Noe JC, Munro CL et al (2005) Identification of virulence determinants for endocarditis in *Streptococcus sanguinis* by signature-tagged mutagenesis. *Infect Immun* 73:6064–6074
7. Xu P, Alves JM, Kitten T, Brown A, Chen Z et al (2007) Genome of the opportunistic pathogen *Streptococcus sanguinis*. *J Bacteriol* 189:3166–3175
8. Ge X, Xu P (2012) Genome-wide gene deletions in *Streptococcus sanguinis* by high throughput PCR. *J Vis Exp* 69:e4356
9. Gaustad P, Havardstein LS (1997) Competence-pheromone in *Streptococcus sanguis*. Identification of the competence gene comC and the competence pheromone. *Adv Exp Med Biol* 418:1019–1021
10. Das S, Kanamoto T, Ge X, Xu P, Unoki T et al (2009) Contribution of lipoproteins and lipoprotein processing to endocarditis virulence in *Streptococcus sanguinis*. *J Bacteriol* 191:4166–4179
11. Turner LS, Das S, Kanamoto T, Munro CL, Kitten T (2009) Development of genetic tools for in vivo virulence analysis of *Streptococcus sanguinis*. *Microbiology* 155:2573–2582
12. De Berardinis V, Vallenet D, Castelli V, Besnard M, Pinet A et al (2008) A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Mol Syst Biol* 4:174

Gene Essentiality

Methods and Protocols

Lu, L.J. (Ed.)

2015, XI, 248 p. 42 illus., 15 illus. in color., Hardcover

ISBN: 978-1-4939-2397-7

A product of Humana Press