Rapid Automated Characterization of Transposon Insertion Mutants in *Desulfovibrio vulgaris* Hildenborough by srnPCR

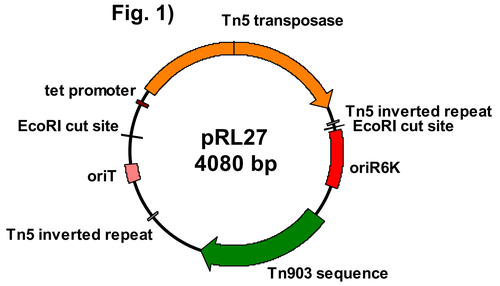
Matthew D. Shirley, Joseph A. Ringbauer, Judy D. Wall
Biochemistry Department, University of Missouri-Columbia

Introduction

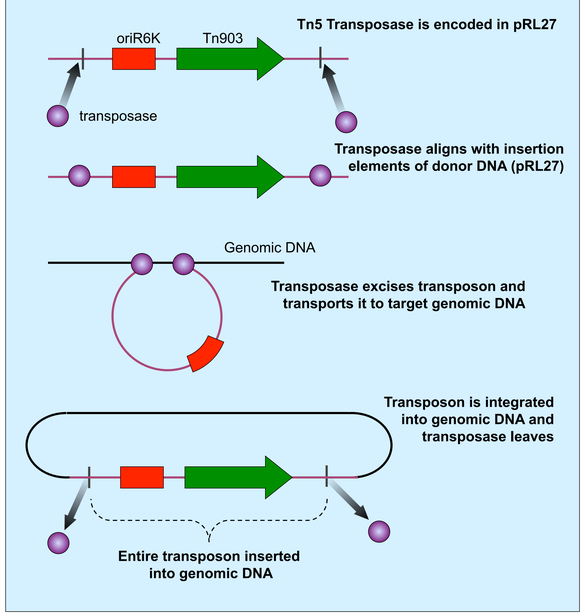
Tn*5* transposon mutagenesis occurs by a mechanism in which a segment of DNA (transposon) encoded in a plasmid is inserted into genomic DNA (the target) by a conservative (cut-and-paste) mechanism *(Fig. 2)*. When the insertion position is in a coding sequence or regulatory region of DNA, the insertion results in a mutation. The plasmid pRL27 was used to transform *Desulfovibrio vulgaris* Hildenborough by electroporation. Transposon insertion mutants were identified by their ability to grow in the presence of kanamycin. To locate the insertion site of the transposon, in theory, one should be able to directly sequence from the transposon into chromosomal DNA *(Fig. 3.1)* and identify the mutation site by comparison with the known genome using BLAST. Unlike sequencing of plasmid DNA or PCR products, direct genomic sequencing has a limited success rate. Therefore, a method of enriching the transposon-flanking sequence is needed. Nested semi-random PCR *(Fig. 3.2)* is an efficient and cost effective enrichment method. Sequencing these enriched products allows us to identify the transposon insertion site. The factors that influence characterization success rate are: frequency and location of priming sites, reaction volume, and reaction conditions (annealing temperature, extension time, etc.). By varying these factors, we have developed an efficient and reliable method for characterizing transposon insertion mutants. Utilizing high-throughput robotics and nested semi-random PCR, we have identified single gene mutants that may provide valuable biological data.

***Our objective** is to develop an rapid, reliable method for locating where transposon insertion sites have occurred in each mutant.*

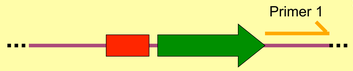
Fig. 1)



2) Transposon Insertion Mechanism



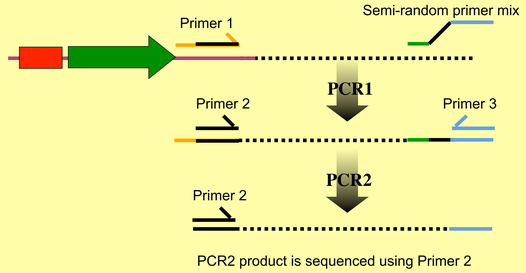
3.1) Direct Genomic Sequencing



Sequencing with primer internal to transposon insertion gives flanking DNA sequence.

Disadvantages:
- *Low primer binding specificity. D. vulgaris* genome = 3,570,858bp. Specific priming sequence = 20bp.
- *Secondary Structures.* Hairpins and supercoiled DNA are more difficult for enzyme to access.

3.2) Nested Semi-Random PCR



PCR2 product is sequenced using Primer 2

Advantages:
+ More enriched sequencing template = higher sequencing success rate
+ Less overall cost (fewer failed reactions)

Semi-random primer mix (PCR1):
CEKG2A: GGC CAC GCG TCG ACT AGT ACN NNN NNN NNN AGA G
CEKG2B: GGC CAC GCG TCG ACT AGT ACN NNN NNN NNN ACG CC
CEKG2C: GGC CAC GCG TCG ACT AGT ACN NNN NNN NNN GAT AT
CEKG2D: GGC CAC GCG TCG ACT AGT ACN NNN NNN NNN AAC GC
CEKG2E: GGC CAC GCG TCG ACT AGT ACN NNN NNN NNN TCG AC

Primer 1 (PCR1):
tpnRL17-1: AAC AAG CCA GGG ATG TAA CG

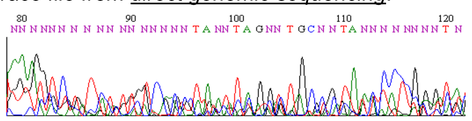
Primer 2 (PCR2):
tpnRL17-2: GGC CAC GCG TCG ACT AGT AC

Primer 3 (PCR2):
CEKG4: GGC CAC GCG TCG ACT AGT AC

4) Comparison of Sequencing Results

Trace file from direct genomic sequencing:

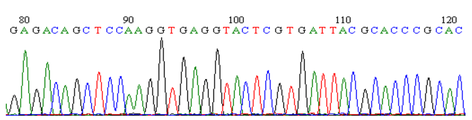
A.



- Very little separation between peaks
- "Background" peaks caused by low primer specificity

Trace file from sequencing nested PCR products:

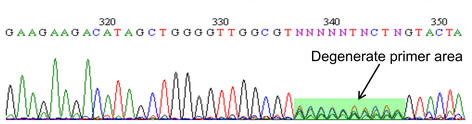
B.



+ Good separation between peaks
+ Better signal-to-noise ratio
+ Bases miscalled by software can sometimes be "called" by hand

Trace file showing the degenerate primer region:

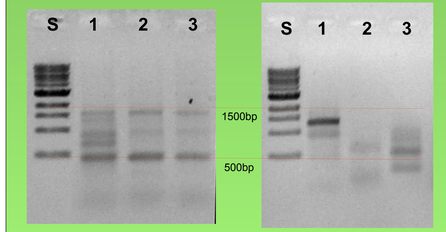
C.



* Sequencing usually terminates after the primer sequence is encountered
* Larger PCR products give more usable sequence before terminating
* A ~1kb product is ideal

5) Analysis of Nested PCR Products

PCR1: Primer 1 + CEKG2 PCR2: Primer 2 + CEKG4

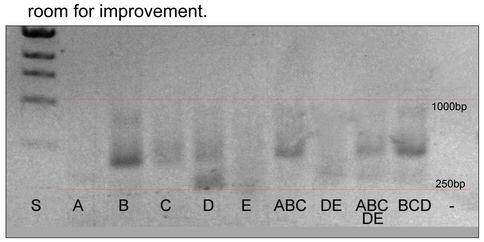


PCR1 Template:
1. Tn903 Mutant JW3068
2. Tn903 Mutant JW3069
3. Tn903 Mutant JW3071
S: 1kb DNA ladder

PCR2 Template:
1. PCR1 JW3068
2. PCR1 JW3069
3. PCR1 JW3071
S: 1kb DNA ladder

6) Semi-Random Primer Design

- ✓ Size of *D. vulgaris* genomic sequence: 3,570,858bp
- ✓ Target PCR product size is 1kb.
- ✓ An ideal primer should anneal every 3,570 bp.
- ✓ Designed to anneal sufficiently close to the transposon to yield a product.
- ✓ Theoretically a mix of three primers (A, B, & C) may cover the entire genome in 1kb or less increments.
- ✓ *D. vulgaris* mean GC content: 63%
- ✓ Published primers have GC content of 20%, 50%, and 80%, and anneal at 1404, 6846, and 7126 unique sites.
- ✓ These primers were optimized for yeast, so there is some room for improvement.

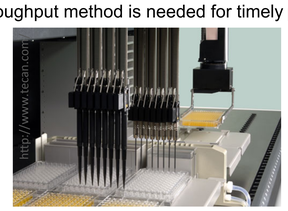


Above: Agarose gel electrophoresis of PCR2 products from JW3101 mutant. Different combinations of degenerate primers were used, including newly designed D & E.

- ✓ Designed two new primers (D & E) with 60% GC, that anneal at 3681 and 5226 unique sites.
- ✓ Semi-random primer amplification is not an exact process; different transposon mutant templates will be differentially successful with the current set of primers.

High-Throughput Methods

- ➢ *D. vulgaris* genome is predicted to contain 3634 protein-coding genes.
- ➢ 5,000 single transposon colonies were isolated and picked into 96-well plates.
- ➢ A high throughput method is needed for timely project completion.



- ➢ Through a collaboration with the MU DNA Core Facility, liquid handling robotics will expedite PCR amplification and sequencing.

BLAST as a Tool For Location

- **B**asic **L**ocal **A**lignment **S**earch **T**ool <http://www.ncbi.nlm.nih.gov/BLAST/>
- Compares sequence data to the known genomic sequence
- Identify what sequence and gene is interrupted
- Allows characterization of genotype and hypothesis of phenotype

Mutants Recently Characterized

Gene	Description	DVU#	Stress	Time	Log[R]*
	transcriptional regulator, putative	2527	50C	15min	1.22
	hypothetical protein	2135	pH10	120min	1.47
purF	amidophosphoribosyltransferase	161	Air	240min	3.63
	thioesterase family protein	159	50C	15min	2.83
dcm	type II DNA modification methyltransferase, putative	1515	Nitrate	240min	-1.13
	lipoprotein, putative	1758	Air	240min	-3.10
	methyl-accepting chemotaxis protein	1869	Air	240min	-1.32
pstB-1	phosphate ABC transporter, ATP-binding protein	1084	50C	15min	-1.73
	hypothetical protein	1750	50C	15min	1.35
	conserved hypothetical protein	273	NaCl	120min	1.62
	hypothetical protein	2690	Air	120min	1.62
clpA	ATP-dependent Clp protease, ATP-binding subunit ClpA	1602	Air	120min	2.93
	sensory box/GGDEF domain/EAL domain protein	3064	Air	120min	1.17
	outer membrane efflux protein	3097	pH10	120min	-1.19
	glycosyl hydrolase, family 3	2239	NaCl	120min	-1.76
	RND efflux system, outer membrane protein, NodT family	62	NaCl	120min	1.65
	tail protein, putative	2728	50C	15min	2.11
	membrane protein, putative	308	NaCl	120min	2.75
	GGDEF domain protein	3106	Air	240min	1.77
	peptidase, PfpI family	1933	50C	15min	-1.22
motA-1	chemotaxis protein MotA	50	50C	15min	1.16
	YjeF-related protein	1910	Air	240min	-1.38
glnD	protein-P-II uridylyltransferase, putative	1233	:	:	:
cysD	phosphoadenosine phosphosulfate reductase, putative	1566	Air	240min	2.36
	Sua5/YciO/YrdC/YwlC family protein	1255	:	:	:
mltA	transglycosylase, putative	887	:	:	:
hisI	phosphoribosyl-AMP cyclohydrolase	113	50C	15min	-1.68
zraS	sensor protein ZraS	3382	50C	15min	2.51

* Log[R] is the measure of regulation under stress relative to normal gene expression. Positive/negative = increased/decreased expression. Larger number = more/less expression.

Summary

- Direct genomic sequencing gives poor quality sequence data
- Nested semi-random PCR enriches the sequence flanking transposon insertion sites
- Sequence data from nested PCR products is more accurate than direct genomic sequencing
- High throughput robotics will accelerate transposon mutant characterization in *D. vulgaris.*

References:

Kristin T. Chun, et al. 1997. "Rapid Amplification of Uncharacterized Transposon-tagged DNA Sequences from Genomic DNA." Yeast **13**(3): 233-240.

Martha M. Howe, D. B. (1989). Mobile DNA. Washington, DC, American Society for Microbiology.

Rachel, A. L., M. W. Marlena, et al. (2002). "Genetic analysis of pigment biosynthesis in Xanthobacter autotrophicus Py2 using a new, highly efficient transposon mutagenesis system that is functional in a wide variety of bacteria." Archives of Microbiology **V178**(3): 193.