

GENE AMPLIFICATION IN *ACINETOBACTER*

by

ANDREW B. REAMS

(Under the Direction of Ellen L. Neidle)

ABSTRACT

This dissertation reports the discovery of gene amplification in the soil bacterium *Acinetobacter* sp. strain ADP1. Amplification mutants were selected from a strain that lacks two transcriptional activators. Without these regulators, it does not express two *cat*-gene transcriptional units at levels high enough to grow on benzoate. Mutant colonies were selected on benzoate medium at a frequency of 10^{-8} within a three-week period. Of 105 independently isolated mutants, 104 carried tandem head-to-tail repeats of a chromosomal segment, or amplicon, containing the *cat* genes. The amplicons ranged in size from approximately 12 to 290 kbp and varied in copy number from 3 to more than 30. Gene amplification occurred via a two-step process whereby an initial duplication event was followed by the generation of higher amplification. Recombinational sites involved in the initial duplication event were isolated using a transformation assay that takes advantage of the natural transformability of these *Acinetobacter* strains. Analysis of sequences from 72 mutants revealed that most of the initial duplications were generated by illegitimate recombination. In some cases, the recombination events were site specific in a fashion that was independent of short homologous repeats. These duplications occurred independently of *recA* and spontaneously in the absence of selection. In contrast, the second amplification step required RecA and was only apparent after selection. Additional

experiments addressed the issue of the delay in amplification mutant colony formation during the original selection period. These studies tested the possibility of “adaptive amplification”.

According to this controversial theory, amplification rates can increase in response to selective conditions via inducible mechanisms for surviving stress. However, our results indicated that this delay in colony formation resulted from a natural selection process whereby higher amplification from preexisting duplications was slowly generated through a series of homologous recombination events. Thus, these mutants did not arise adaptively. Furthermore, this system presented new evidence that supports the theory that the co-amplification of functionally related genes is a driving force in the evolution of higher-order gene clusters. This new model system for studying gene amplification illuminates a process with important implications for evolution, drug resistance, increased virulence, bioremediation, and cancer.

INDEX WORDS: Gene amplification, Tandem duplication, Gene cluster, Co-amplification, Novel junction, Illegitimate recombination, RecA, *Acinetobacter*, ADP1, Bacteria, β -ketoadipate pathway, Benzoate, Adaptive, Bioremediation, Cancer

GENE AMPLIFICATION IN *ACINETOBACTER*

by

ANDREW B. REAMS

B.S., California State University - Chico, 1995

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

© 2003

Andrew B. Reams

All Rights Reserved

GENE AMPLIFICATION IN *ACINETOBACTER*

by

ANDREW B. REAMS

Major Professor:	Ellen Neidle
Committee:	Timothy Hoover Anna Karls Mark Schell Lawrence Shimkets

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2003

DEDICATION

I dedicate this work to my family for their love and support throughout my life. I also dedicate this to my wife, Lisa, who has been incredibly supportive during my graduate school career.

“By endurance we conquer”

- Sir Ernest Henry Shackleton

ACKNOWLEDGEMENTS

First, I would like to thank my mentor Ellen Neidle for her generosity and her guidance. Under her leadership, I was given the freedom to experiment with my ideas, while being encouraged to question my results. Thanks to other members of the Neidle family: Matt Eby for his friendship and stimulating dialogue, Lauren and Becky for their guidance during my first years as a graduate student, and Sandra, Todd, and Obidi for their camaraderie and friendship. A sincere thanks to my committee members, Tim Hoover, Mark Schell, Anna Karls, and Larry Shimkets, for their time and advice. Also, to my friends from Athens, I hope to continue our friendships into the future.

I would especially like to thank Jeff Dowding for inspiring me to become a scientist. You have provided me with invaluable advice and admirable friendship. To all my friends from Mission Viejo who have always been there for me. I am very fortunate to have such a neighborhood.

Thank you family for your unconditional support. You are the source of my strength and stability. I would also like to acknowledge my recently acquired in-laws for their encouragement and support. To my wife, Lisa, who I owe all my academic honors. You have made me realize the reason there are more important things in life than just work.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
CHAPTER	
1 GENE AMPLIFICATION: USING A NOVEL MODEL SYSTEM TO ADDRESS QUESTIONS OF BROAD BIOLOGICAL SIGNIFICANCE	1
Introduction	1
Review: Selection for Gene Clustering by Tandem Duplication	3
Conclusion.....	24
2 GENOME PLASTICITY IN <i>ACINETOBACTER</i> : NEW DEGRADATIVE CAPABILITIES ACQUIRED BY THE SPONTANEOUS AMPLIFICATION OF LARGE CHROMOSOMAL SEGMENTS	26
Summary	27
Introduction	28
Results	31
Discussion	49
Experimental Procedures.....	56
Acknowledgements	62
3 GENE AMPLIFICATION INVOLVES SITE-SPECIFIC SHORT HOMOLOGY- INDEPENDENT ILLEGITIMATE RECOMBINATION IN <i>ACINETOBACTER</i> SP. STRAIN ADP1	63

Summary	64
Introduction	65
Results	68
Discussion	86
Experimental Procedures.....	94
Acknowledgements	99
4 GENE AMPIFICATION IN <i>ACINETOBACTER</i> : EVALUATION OF ADAPTIVE MUTABILITY	100
Summary	101
Introduction	101
Results and Discussion.....	104
Experimental Procedures.....	127
5 CONCLUSION.....	132
REFERENCES	135

CHAPTER 1

GENE AMPLIFICATION: USING A NOVEL MODEL SYSTEM TO ADDRESS QUESTIONS OF BROAD BIOLOGICAL SIGNIFICANCE

Introduction:

During my Ph.D. research I discovered gene amplification in the soil bacterium *Acinetobacter* sp. strain ADP1. Gene amplification underlies many physiologically significant biological processes such as drug resistance and increased virulence in bacteria. In recent years, considerable interest in gene amplification has developed due to the discovery that in higher eukaryotes gene amplification is a major route to oncogenesis and cancer progression. Still, the underlying mechanism(s) remain unclear for any organism. The biological significance of gene amplification amongst diverse organisms and the realization that several *Acinetobacter* characteristics make it an ideal experimental system, resulted in focusing on this topic for my doctoral dissertation. One of my goals was to utilize ADP1 as a model system to understand gene amplification and to elucidate the mechanism(s) behind this important phenomenon. This model system may provide valuable insights for understanding some aspects of pathogenesis as well as for understanding gene amplification in higher cells.

The following chapters describe various aspects of this model system. The second chapter, which was recently published in *Molecular Microbiology*, reports the original discovery of gene amplification in strain ADP1 and describes the methods used for genomic analyses of mutants. This chapter also describes a transformation assay, based on the natural competence of

these *Acinetobacter* strains, which enables the precise amplicon endpoints to be determined. This assay is very powerful because it allows the characterization of distant chromosomal regions involved in the duplication recombination events.

The third chapter, which was recently submitted for publication, reports the size of the amplified segment and the level of amplification in over 100 independently isolated mutants. By taking advantage of the transformation assay, the initial duplication event was elucidated by the sequence analysis of over 70 amplicon boundary points. This large-scale study suggests that most duplications are generated by illegitimate recombination that can involve site-specificity. This type of site specific recombination has not been previously characterized and suggests a novel mechanism.

The final research chapter describes studies testing whether gene amplification is induced during selection. The possibility that amplification rates are augmented by stress-induced mechanisms has been proposed by the controversial theory of “adaptive amplification”. My results demonstrate that tandem duplications arise independently of the presence of selection. Further experiments show that duplication mutants may take several weeks to generate higher amplification necessary to form a colony under selection. These results suggest amplification arose through natural selection rather than being induced by an “adaptive” mechanism.

These gene amplification studies have additional evolutionary significance. The selection system yields amplification mutants nearly exclusively. This result depends on the requirement for the co-amplification of two different transcriptional units. The genes in these units are part of a supraoperonic cluster of functionally-related genes involved in aromatic compound catabolism. These clusters extend beyond the operon level and the question of how such clusters have evolved has received increasing attention. Recently, the analysis of whole

genomes has shown that genes are not randomly dispersed on chromosomes but rather evolutionary forces appear to select an organization that reflects function. Horizontal gene transfer has been proposed as one factor driving the clustering of functionally-related genes. My studies support the additional possibility that gene amplification is a driving force in how such clusters have evolved. Consequently, Dr. Ellen L. Neidle and I were invited to write a review article discussing the evolutionary significance of gene amplification in the formation and maintenance of gene clusters. The next portion of this introductory chapter presents a draft of this review article.

Review: Selection for Gene Clustering by Tandem Duplication

Higher-order Gene Clustering

Rearrangements constantly re-sculpt genomic landscapes. Yet, genome sequencing and molecular biology have revealed that chromosomes are significantly more organized than ever imagined (Lawrence, 2003). One important form of organization is the clustering of functionally-related genes. Gene clusters arise and are maintained in both prokaryotic and eukaryotic genomes (Lee and Sonnhammer, 2003). In general, the degree of clustering is greater in prokaryotes than in eukaryotes. These clusters can be located either chromosomally or extrachromosomally and can vary in size from 2 genes up to 90 genes (Jordan *et al.*, 2001). In prokaryotes, genes can also be arranged within operons, allowing for transcriptional co-regulation (Jacob *et al.*, 1960, Jacob and Monod, 1962). Similar systems for co-expression may exist in some eukaryotic gene clusters (reviewed in Lawrence, 2003). However, it has been argued that transcriptional co-regulation does not explain the clustering of genes, since genes widely scattered on the chromosome also have the ability to be co-regulated. Furthermore, some

gene clusters extend beyond the operon level and contain multiple transcriptional units. Such clusters are sometimes referred to as higher-order gene clusters.

Genes in a particular cluster usually encode proteins that serve a common function or contribute to a joint phenotype. Such clusters are termed supraoperonic clusters (Rosenberg and Hegeman, 1969) or gene islands (Hacker and Kaper, 2000). Examples of these higher-order clusters include pathogenicity islands, catabolic islands, nitrogen fixation gene clusters, membrane bound protein clusters, photosynthetic gene clusters, and clusters involved in specialized functions such as magnetic behavior. These clusters are believed to reflect genes valuable for adapting to selective pressures in particular types of environments.

Operons in bacterial genomes are poorly conserved over long periods of time. The gene order, gene content, and regulatory mechanisms of operons can vary greatly, even in closely related species. However, this rearrangement tends to be a conservative process, whereby the rearrangements between and within operons often maintain individual genes in very specific functional and regulatory contexts. Conservatively-rearranged operons have been termed “uber-operons” (Lathe *et al.*, 2000).

These conservative rearrangements appear to also apply to higher-order gene clusters. One example of these rearrangements is illustrated by the gene order shuffling in aromatic compound degradation genes associated with the β -ketoadipate pathway. Both fungi and bacteria use this pathway for growth with a diverse range of aromatic compounds (Harwood and Parales, 1996). Genes associated with this pathway are usually clustered. These clusters can be quite large, extending to more than 40 genes, and include multiple transcriptional units (Parke and Ornston, 2003; Smith *et al.*, 2003). When comparing these gene clusters amongst other organisms there is a striking degree of gene order shuffling (Young *et al.*, 2001; Buchan *et al.*,

2003). Thus, the existence of higher-order clusters throughout genomes and the conservation of functionally-related clusters clearly indicate this level of organization is important. However, the evolutionary forces leading to the origin and conservation of higher-order clusters are not yet fully understood.

Several theories concerning the formation and maintenance of multiple transcriptional unit clusters have been proposed (Lawrence, 2003). Two models are the “selfish operon model” (Lawrence and Roth, 1996) and the “repair model” (Hartman *et al.*, 1960). For both models, the advantage of clustering increases as the genes are brought closer together. The selfish operon model states that clustered genes allow co-transfer of genes encoding a selectable phenotype to a new host (Roth *et al.*, 1996). The clustered genes are selfish in that the genes themselves are preserved instead of the original host. Several examples of clusters that have undergone horizontal gene transfer support this theory (Lawrence, 2003; Davis and Waldor, 2003; Cotter and DiRita, 2000).

The repair model proposes that clustering provides a means for the simultaneous repair of multiple defective genes. For example, a particular bacterial phenotype may only rarely be placed under selective conditions. During the long periods of relaxed selective pressure, the genes contributing to a particular function can accumulate mutations and the number of mutant cell clones might become significantly large. These mutations can cause the function to be lost and can be reacquired only by mating with a cell which can provide a new version of the genes. The repair model proposes that clustering of functionally-related genes provides a single fragment as a template for simultaneously repairing these multiple mutated genes, presumably by gene conversion (Hartman *et al.*, 1960).

Another theory proposes that the clustering of functionally related genes may contribute a selective advantage to an organism by enabling the set of related genes to be co-amplified by tandem duplication (Roth *et al.*, 1996). We refer to this theory as the “co-amplification model”. This model involves the use of tandem duplication as a primitive regulation mechanism for increased gene expression. By increasing the gene dosage through gene amplification, a concurrent increase in the expression of genes encompassed on the amplified segment can occur and be selectively advantageous under some conditions. If the genes were scattered, multiple simultaneous mutations might be required to achieve the same effect, and therefore the advantageous outcome would be much less likely to occur. Furthermore, through co-amplification each gene increases in equal copy number and results in properly balanced expression of the entire gene set. Keeping the relative expression levels of the genes in balance may be important in a common metabolic pathway to avoid the accumulation of intermediates that can cause inhibitory effects. Additionally, since the amplification process is reversible, the cells can revert to the original non-mutated state once relieved of selective pressure. The co-amplification model is similar to the selfish operon and repair models in that the greater the proximity of genes, the more likely that they can confer a selective advantage. If a block of functionally-related genes is frequently amplified under stress, a selective force for clustering is provided. The three different theories for the selection of gene clusters are not mutually exclusive. Rather, they may all contribute to the evolution of genomic arrangements that maintain functionally related genes in proximity to each other. As described below, experimental evidence demonstrates that the co-amplification of genes that are functionally related but not cotranscribed can be readily selected.

Gene Amplification

Gene amplification is the increase in copy number of a DNA segment, or amplicon (Roth et al., 1996). This type of chromosomal rearrangement may allow a clone to overcome its inability to proliferate under growth-limiting conditions by augmenting the expression of a particular gene on the amplicon. Some variant cells have been shown to carry long tandem arrays of gene copies. Once tandem arrays of long homologous sequences are generated, recombination between the repeat sequences causes amplification to be highly reversible. In fact, deamplification may occur rapidly after the removal of a selection pressure or the appearance of a permanent mutation which alleviates the need for amplification. For this reason, amplification may be elusive to identify and difficult to study.

Nevertheless, gene amplification has been well documented as a means for adaptation to environment stress by increasing protein dosage without permanently changing the genome (Romero and Palacios, 1997). In microbes, it contributes to such adaptations as antibiotic resistance, resistance to heavy metals, growth on exotic carbon sources, and increased virulence. In plants, gene amplification has been correlated with herbicide resistance (Harms et al., 1992, Shyr *et al.*, 1992). In humans and other mammals, amplification is not known to occur in normal development; however, it is observed readily in virtually all types of cancer cells, in which amplification of cellular oncogenes contributes in important ways to tumorigenesis and metastasis (Stark, 1993).

In most reports of gene amplification, the selection targets the increase in copy number of a single gene. For example, the amplification of a single gene encoding antibiotic resistance may be selected in medium containing elevated antibiotic concentrations. However, the amplicon very rarely encompasses only the targeted gene. Instead, the amplicon is usually relatively large

and includes several neighboring genes which are presumably not targeted by the selection. The co-amplification of non-targeted genes may have lethal effects or lead to non-lethal side-effects such as increased mutagenesis. In contrast, examples exist in which the co-amplification of multiple transcriptional units provides a selective advantage for the cell. For example as described in this thesis, in *Acinetobacter* sp. strain ACN293, selective growth on medium containing benzoate as the sole carbon source requires increased expression of two clustered transcriptional units, the *catA* gene and the *catBCIJFD* operon. This selection depends on the co-amplification of both transcriptional units. Deamplification, accompanied by a decrease in the ability to grow under selective conditions, was seen upon culturing under non-selective conditions (Reams and Neidle, 2003). This example clearly illustrates the advantage of having clustered genes by demonstrating the ability of an organism to co-amplify in certain environments without a requisite permanent chromosomal mutation.

In prokaryotes, tandem gene amplification generally is presumed to occur via a two-step mechanism in which an initial duplication event is followed by the subsequent generation of higher amplification (Roth *et al.*, 1996; Romero and Palacios, 1997). The formation of the initial tandem duplication does not cause the strain to lose expression of any genetic information as long as the duplication is not entirely within a single functional genetic element or transcriptional unit (gene or operon). Tandem gene amplification usually results in the head-to-tail amplification of an amplicon that is uniform in size. The size of the amplicon is determined by the initial tandem duplication event, which is believed to occur by a sister strand unequal exchange. As a result of this recombination event a novel junction sequence is generated by the fusing of two normally distant regions determining the amplicon endpoints. Tandem duplication

formation results in long sequence repeats which can support RecA-dependent recombination events that increase the copy number of the amplicon.

The initial tandem duplication event can occur at a high frequency independently of the presence of selection. The duplication frequency for a specific locus varies widely, from 10^{-2} to 10^{-5} (Romero and Palacios, 1997). In an unselected population of *Salmonella*, Roth *et al.* estimated that the typical locus is duplicated once in every 1,000 cells. It is estimated that nearly 10% of the cells in these cultures carry a duplication of some region of the chromosome. Duplications may affect almost any locus on the bacterial chromosome. Sites restricted from duplication may be those interfering with segregation or involving a gene or genes that when present in duplicate inhibit growth (Roth *et al.*, 1996).

Tandem duplications vary greatly in size ranging from a few base pairs to several megabases (Lupski *et al.*, 1996). In *Salmonella enterica* and *Escherichia coli*, duplications range up to about half of the chromosome (Heath and Weinstock, 1991; Anderson and Roth, 1981; Roth *et al.*, 1996). Cells carrying highly amplified genes may have some amplicon size restriction due to the constraints of containing a large number of long tandem repeats. However, amplicons greater than 100 kilobase pairs in size have been reported to have amplified over 5-fold (Ravatn *et al.*, 1998; Reams and Neidle, 2003). Such large-size amplicons invariably contain multiple transcriptional units.

Evaluation of the entire *S. enterica* chromosome for duplication occurrence revealed that identical segments are commonly duplicated in tandem (Anderson and Roth, 1981). These “hotspot” duplication zones, with frequencies of occurrence as high as 3%, have been described in every bacterial system in which tandem duplication recombination sites have been analyzed, whether chromosomal or plasmid-borne. A common feature in every characterized hotspot is the

presence of repeated sequences larger than 1 kbp in direct orientation which flank the region undergoing tandem duplication. Homologous recombination between the repeated sequences may lead to either a tandem duplication or to a deletion of the intervening sequences. These rearrangements are thought to be formed by mechanisms involving unequal exchanges or circle-excision and reinsertion (Romero and Palacios, 1997).

The specific nature of the repeated element is unimportant in generating a tandem duplication. There are common inhabitants of many genomes among these repeated sequences that may participate in generating tandem duplications, such as repeated rRNA (*rrn*) operons and IS elements of different kinds. Alternatively, species-specific repeated sequences can participate in duplication formation, such as the Rhs elements found only in *E. coli* (Lin *et al.*, 1984), or the multiple copy nitrogenase operons in *Rhizobium etli* (Romero and Palacios, 1997). The formation of duplications is highly dependent on RecA function when long recombining sequences are used as recombination substrates (Roth *et al.*, 1996).

Mobile Elements: Integration accompanied by amplification

Large direct repeats commonly are used in generating tandem duplications. This observation may suggest a linkage between mobile genetic elements and the ability to amplify. As noted in the selfish operon model, clustered genes are commonly located on mobile elements (Lawrence, 2003). In some of these mobile elements, the intervening genes are flanked by large direct repeat sequences, such as insertional sequences (IS). These IS sequences may play multiple roles. First, they may be involved in the transposition of the element. Also, the IS sequences may provide substrates for homologous recombination in the generation of tandem duplications (Peterson and Rownd, 1983). Tandem duplication and higher amplification of a newly

integrated element may provide a selective advantage to the new host by increasing gene expression of a selectable phenotype encoded on the mobile element. Indeed, there are several examples in which integration of a genetic element is accompanied by amplification of the element. For example, reports exist of amplification of antibiotic-resistance determinants of plasmid origin when they are forced to integrate into the chromosomes of new hosts such as *E. coli* (Guttersen and Koshland, 1983), *Streptococcus pneumoniae* (Vasseghi and Claverys, 1983) and *Bacillus subtilis* (Young, 1983).

Other examples in which integration is accompanied by amplification have been reported in several cases involving clustered genes providing the ability to degrade chloroaromatic compounds. Several bacterial species, including certain *Pseudomonas* species, contain the genes to degrade catechol but not chlorocatechol. Investigators have attempted to widen the substrate-specificity of these organisms by integrating into the chromosomes gene clusters for chlorocatechol degradation derived from other species. In *Pseudomonas* sp. strain B13, the *clc* element is a conjugative transposon approximately 105 kb in size that encodes the capacity to metabolize 3-chlorobenzoate, 3-chlorocatechol, 5-chlorocatechol, and possibly aminophenol (van der Meer and Sentchilo, 2003). The genes for chlorocatechol degradation are encoded by the *clcR-clcABDE* gene cluster located on the *clc* element. The *clc* element is able to transfer itself to several different γ - and β -Proteobacteria under a variety of environmental conditions (Muller *et al.*, 2003). Selection for transduction by growth on chlorocatechols yields cells with an integrated *clc* element accompanied by tandem amplification. In *Pseudomonas putida* F1, the integrated 105 kbp *clc* element was amplified six-fold to increase the genome size by at least 10%. Flanking insertional sequences seemed to be the targets for recombination events to create tandem duplications. Once formed, the amplified structures were unstable and they deamplified

when grown under nonselective conditions (Ravatn *et al.*, 1998). In a similar case, the *tcbR-tcbCDEF* gene cluster coding for chlorocatechol degradation was obtained from *Pseudomonas* sp strain P51 and cloned into a minitransposon. Integration accompanied by tandem amplification of these genes was observed in *Pseudomonas* sp. strain KT2442 cells, providing the ability to grow on chlorocatechols (Klemba *et al.*, 2000). These two cases demonstrate that the integration of mobile elements carrying functionally related clustered genes may be accompanied by co-amplification of these genes to allow for adaptation to degrade chlorocatechol.

Sometimes amplification of gene clusters involved in chlorocatechol degradation does not involve integration. For example, the duplication of the *cbnR-cbnABCD* gene cluster from *Alcaligenes eutrophus* NH9 was found to be associated with increased fitness following long-term growth in liquid batch cultures of chlorobenzoate (Ogawa and Miyashita, 1995; Ogawa and Miyashita, 1999). Additionally, a duplication of *clcRA* from the *clcRABD* locus in chlorobenzene-degrading *Ralstonia* sp. strain JS705 was observed (van der Meer *et al.*, 1998). This evidence suggests that amplification of clustered genes has been important in the adaptation of bacteria to chlorocatechol degradation in environments contaminated with chloroaromatics (Ghosal and You, 1988; Rangnekar, 1988; Klemba *et al.*, 2000).

Tandem Duplication of Gene Clusters Using Large Direct Repeats

Whether mobile or not, genetic elements resembling the organization of transposable elements often are flanked by direct repeat sequences. While these repeat sequences initially may have served for the transposition event, they may also provide sequences of homology for recombination in the generation of tandem duplications. Furthermore, alternative direct repeat

sequences containing no known insertion sequences may provide sites for tandem duplication formation (Peterson and Rownd, 1983). Large or small homologous regions may be closely or widely spaced and are numerous on bacterial chromosomes (Prozorov, 2001). It has been suggested that the positioning of some direct repeats has evolved specifically for gene duplication (Lin *et al.*, 1984; Roth *et al.*, 1996; Romero and Palacios, 1997). These tandem duplications, by relying on homologous recombination events, would occur at a relative high frequency. When the environmental conditions change, higher amplification of the preexisting tandemly duplicated genes would be selected (Peterson and Rownd, 1985).

A classic example of gene clusters using flanking direct repeats for duplication formation is the amplification of *Vibrio cholerae* toxin genes during intestinal growth in rabbits (Mekalanos, 1983). The genes for cholera toxin biosynthesis reside on a filamentous phage called the CTX element, which is present on the chromosome of toxigenic strains but absent in nontoxigenic strains. The CTX genetic element has the structure of a composite transposon, with a 4.5-kbp central core region flanked by one or more copies of a 2.7 kbp repetitive sequence, RS1 (Waldor and Mekalanos, 1996). The core of the CTX element is known to carry at least six genes, including *ctxAB* (encoding the A and B subunits of cholera toxin), *zot* (encoding zonula occludens toxins), *cep* (encoding core-encoded pilin), *ace* (encoding accessory cholera enterotoxin), and *orfU* (encoding a product of unknown function). Two different isolates of the pandemic *Vibrio cholerae* El Tor strain selected for intestinal growth in rabbits were reported to have undergone RecA-dependent tandem amplification of the CTX element. The RS1 flanking the CTX element appeared at the novel junctions between the amplicons of both amplification strains, and therefore appears to have been used as the recombinational substrate in the initial tandem duplication event (Mekalanos, 1983; Goldberg and Mekalanos, 1986). Increased copy

number of the pathogenicity island containing the cholera toxin biosynthesis genes correlated with the hypertoxinogenic phenotype found in these derivative amplification strains.

The findings that multiple genes involved in toxin biosynthesis are amplified in *Vibrio cholerae* El Tor as a possible means for increasing virulence may suggest similar amplification mechanisms in some other toxin-producing pathogens. Potent exotoxins of bacterial pathogens often are encoded by accessory genetic elements such as bacteriophages, plasmids, and transposons. For example, the genes specifying diphtheria toxin (Freeman, 1951), streptococcal erythrogenic toxin (Zabriskie, 1964), botulinum toxins C and D (Eklund and Poysky, 1974), *E. coli* shiga toxin (O'Brien *et al.*, 1984), and staphylococcal enterotoxin A (Betley and Mekalanos, 1985) all are carried by temperate bacteriophages. In addition, some toxin genes are located on transposable elements or plasmids (Gyles *et al.*, 1974; Mikesell *et al.*, 1983; Hagblom *et al.*, 1985). The similarity between the gene organization of the cholera toxin genes and other toxin genes in pathogens suggests that amplification may be a common mechanism for increasing the expression of such virulence genes during infection. By having large direct repeats flanking these genes, pathogenic microbes may rely on homologous recombination between the repeats in an unequal exchange in order to generate tandem duplications at a high frequency. Selection for amplification of clustered genes involved in toxin biosynthesis may then enhance the virulence of pathogenic microbes. Potential evidence supporting this theory is provided by McDonough and Butterton (1999) in which a spectinomycin resistance cassette was inserted into the Shiga toxin (Stx) operon of *Shigella dysenteriae* 1. Selection for growth in medium containing spectinomycin yielded strains carrying amplification of the Stx operon region. Duplication of the 22.4 kbp putative composite transposon carrying the Stx operon occurred between the flanking IS600 insertional sequences (McDonough and Butterton, 1999). Although there is no

direct evidence that amplification of the Shiga operon leads to increased virulence in *S. dysenteriae* 1, it is clear that higher amounts of Stx lead to increased virulence in Shiga toxin-producing *E. coli* strains (Acheson *et al.*, 1998; Wagner *et al.*, 1999; Ritchie *et al.*, 2003).

These examples suggest that amplification of clustered virulence genes may provide pathogens a means for adaptation during host invasion. In a related example, gene amplification of the capsulation gene cluster in *Haemophilus influenzae* was found in one-third of the strains isolated from blood or cerebrospinal fluid of children with meningitis (Cornet *et al.*, 1994). In these populations, the capsulation gene cluster is found on an 18 kbp DNA segment flanked by direct repeats of an IS-like element. These IS-like sequences apparently were used during the integration of the element in the chromosome, and as substrates for capsulation gene amplification (Kroll *et al.*, 1991). At some point, the 18 kbp segment became fixed in the *H. influenzae* genome as a “stable” tandem duplication, in which the two tandem copies are not identical; one copy has an approximately 1.2 kbp deletion near one end. The duplicated arrangement is believed to serve as a template for further amplification of the capsulation genes (Corn *et al.*, 1993). Increased capsule production caused by gene amplification is thought to be an important factor contributing to virulence (Hoiseth *et al.*, 1986).

The benefits of gene cluster amplification might be extended to other organisms, as evident in symbiotic interactions. One example of this is *Buchnera aphidicola*, a bacterial endosymbiont of aphids. The endosymbiont often provides amino acids to its aphid host. In most free-living bacteria, the chromosomal genes encoding the first enzyme in each biosynthetic pathway of tryptophan (*trpEG*) and leucine (*leuA*) are regulated by feedback inhibition (Pittard, 1996). While it would be advantageous for free-living bacteria to maintain this regulation in order to prevent the overproduction of tryptophan, *B. aphidicola* requires overproduction of

tryptophan and leucine to provide the excess amount to the aphid. One possible means of overcoming feedback inhibition is by amplification of *trpEG* or *leuABCD*. Indeed, the *B. aphidicola* *trpEG* and *leuABCD* gene clusters can be tandemly amplified up to 24-fold (Lai *et al.*, 1994; Thao *et al.*, 1998). This example further demonstrates the selective advantage of having clustered genes. In the case of *trpEG*, these two genes encode the two subunits of anthranilate synthase, the first enzyme of the tryptophan biosynthetic pathway. Co-amplification of these two genes within one segment allows for an equivalent increase in the production of both subunits composing anthranilate synthase. If these two genes were located in different regions of the genome, independent amplification events or another mechanism would be required to overcome inhibition.

The endosymbiont of legumes, *Rhizobium*, exhibits tandem amplification of plasmid-borne gene clusters involved in nodulation and nitrogen fixation. Amplification of these regions is believed to assist symbiosis by enhancing production of nodulation factors and promoting increased nitrogen fixation in the host (Romero and Palacios, 1997). Tandem duplication of these gene clusters frequently occurs by spontaneous recombination events between the large direct repeat sequences generously scattered on the symbiotic plasmid (Flores *et al.*, 2000). The spontaneous duplications created by homologous recombination may facilitate adaptation by allowing amplification of genes important for symbiosis.

Tandem Duplication of Gene Clusters Generated Without Direct Repeats

Duplications also may be formed by recombination between endpoints with very little or no evident homology. These illegitimate recombination events appear to be RecA-independent (Reams and Neidle, unpublished data). In *Acinetobacter* sp. strain ACN293, a selection for

increased expression of two clustered *cat* gene transcriptional units yielded derivative strains carrying tandem co-amplification of both of the transcriptional units (Reams and Neidle, 2003). Analysis of the junction sequences from 70 independent amplification mutants derived from this selection revealed little sequence identity (<8 bp) in the DNA regions that had recombined during the duplication generation (described in chapter three of this thesis). The observance of duplications primarily formed by illegitimate recombination may reflect the restrictions of the selection method. This method required that one recombination site occurs within an approximately 5.5 kb region. Since this requirement may limit the availability of large direct repeats, duplications created by illegitimate recombination could have been preferentially selected.

Several examples have been documented of duplications generated by illegitimate recombination between short sequence repeats (7-15 bp). These examples include regions encompassing the *ampC* or *lacZ* genes in *E. coli* (Edlund *et al.*, 1981; Tlsty *et al.*, 1984; Whoriskey *et al.*, 1987). Duplication generated with or without direct repeats may be important for maintaining gene clusters that have existed in the genome over long periods of time by allowing for their co-amplification. Such amplification may provide a selective advantage under some conditions.

Co-amplification of Non-targeted Genes May Cause Side-Effects

Gene amplification appears a rather crude mechanism as a gene regulatory system because it can cause the co-amplification of genes presumably not favored by the selection that are within close proximity of the targeted genes. This co-amplification of neighboring genes may have lethal effects. For example, duplication of the terminus region is apparently prohibited in *Salmonella*,

possibly because two copies of this region may interfere with chromosome segregation (Roth *et al.*, 1996). Also, the co-amplification of a tetracycline resistance gene is lethal since the overproduction of this gene product is believed to compromise the proton motive force and result in cell death (Hennrickson *et al.*, 2002). Additionally, the co-amplification of neighboring genes may produce non-lethal but detrimental side effects. For example, co-amplification of the gene encoding the error-prone DNA polymerase DinB causes increased mutagenesis (Kofoid *et al.*, 2003). However, tandem amplification is not utilized as a primary means to regulate gene expression. Instead, amplification provides a mechanism to allow adaptation during exposure to unusual or extreme environmental conditions.

Some investigators have attempted to insert genes without a selectable phenotype into regions that are selectively targeted for amplification. For example, co-amplification of a linked gene, *uidA*, was observed in transgenic petunia plant cells selected for amplification of a kanamycin resistance gene, *nptII* (Jones *et al.*, 1994). Co-amplification of genetically-engineered linked genes has been proposed to have applications for increasing the production of various antibiotics (Altenbuchner and Cullum, 1987) and biotherapeutic proteins (Dorai and Moore, 1987; Hendricks *et al.*, 1989).

Importance of the Role of Gene Amplification in Cluster Maintenance

The importance of the role of co-amplification in cluster maintenance may depend on the frequency of exposure of a population to selection. This may be more crucial in organisms such as pathogens and symbionts which normally are exposed to extreme environmental stress (Roth *et al.*, 1996). Another factor may be the duplication frequency of a particular locus. The positioning of flanking direct repeats appears to increase the frequency of duplication formation

of a particular region. Populations carrying a duplication of a targeted region may be selected for higher amplification during environmental stress. However, when dealing with a large number of cells, duplications representing most regions of the genome may be present within the population. The positioning of genes, which when present in multiple copy may cause deleterious or lethal effects, adjacent to the targeted amplified area can influence whether the duplication and amplification of particular regions is advantageous.

Clustered Genes Are More Likely To Co-Amplify

According to the co-amplification model, the ability of a cell to co-amplify a specific group of genes increases with gene proximity. In contrast, genes that are scattered on a chromosome are less likely to be co-amplified. The ability to amplify multiple independent chromosomal segments each covering scattered regions seems unlikely since a single cell would need to contain independent tandem duplications encompassing each of the different genes. Like other mutations, the probability of having both events in a single cell is most likely a product of the probability of each event. Surprisingly, there is at least one documented case in which two different genomic segments are amplified independently. A strain of the protozoan *Leishmania tarentolae* resistant to oxyanion arsenite contains two genes, *gshI* and *pgpA*, amplified on independent segments (Grondin *et al.*, 1997). Concurrent amplification of these genes contributes synergistically to arsenite resistance. *Leishmania* spp. appear unique in their ability to amplify genes, as strains have been reported to spontaneously amplify segments to high copy numbers in the absence of selection (Segovia, 1994).

Divergent Evolution of Gene Clusters Facilitated by Tandem Duplication

For decades, scientists have embraced the idea that gene duplication plays an important role in the establishment of genes with novel functions within homologous gene families (Ohno, 1970). There is also evidence that duplications contribute to the formation of gene clusters. These duplications can provide a means for cluster formation either by increasing the number of genes within a cluster or by duplicating chromosomal segments, leading to the origin of entire novel sets of clustered genes.

Genes within entire clusters can be derived from a single ancestral gene. This is evident in some eukaryotes, in which the α - and β -globin gene clusters and many hox gene clusters arose by duplication and subsequent divergence. Molecular phylogenetic analysis indicates that each of these gene clusters derived from a single ancestral gene that underwent multiple duplication events to form its own cluster (Doolittle, 1987; Finnerty and Martindale, 1999; Pollard and Holland, 2000). However, these clusters appear to be the exception rather than the rule. Most bacterial gene clusters encode proteins catalyzing mechanistically diverse reactions which in union serve a common function or contribute to a phenotype. In general, sequence comparisons suggest that most proteins in a single pathway are unrelated and belong to disparate functional families (Roth *et al.*, 1996).

One theory for the formation of clustered genes is that genes were originally scattered about the chromosome and brought together to form clusters through rearrangements (Roth *et al.*, 1996). These rearrangements include duplications. Prior to duplication, deletions between functionally related genes leading to their clustering likely would remove important genes. However, subsequent to duplication, deletions of redundant nonessential genes or between genes on the duplicated segment would not be detrimental because no genetic information would be

lost. Rearrangements causing clustering of functionally-related genes would be selectively favored. This is similar to the theory proposed for cluster formation by the selfish operon model, in which mutations and rearrangements on the transferred fragment may be tolerated because the genes are not under selective pressure (Lawrence, 2003). Thus, both the duplication model described above and the selfish operon model may explain the formation of clusters extending beyond the operon level. Tandem duplications may contribute further to cluster formation with the placement of two normally distant regions adjacent to each other by nature of the tandem duplication's novel junction sequence. This has been demonstrated by several selections in which genes inhibited in expression are placed under control of alternative promoters by alignment with the novel junction (reviewed in Anderson and Roth, 1977).

Duplications also contribute to paralogous gene cluster formation by copying segments encompassing entire clusters. One might anticipate that genomic evidence for segmental duplication of gene clusters would be quite difficult to obtain. Evidence shows that the majority of new genes that arise by duplication acquire mutations that inactivates them so that they become pseudogenes (Wagner, 2001). Unless the duplication was recent, pseudogenes will diverge in sequence over time and the duplication event will eventually no longer be detectable in the DNA sequence. However, those genes that are retained, because their duplicated functions are useful to the organism or because they have evolved new function, should be identifiable. To determine whether or not gene clusters arose by segmental duplication, sequence identity between the gene clusters in the same species is compared. For these duplications to be signaled it is necessary to locate duplicated sets of genes which have the same order in both sets. To what extent these duplicated sets are still visible in the genome depends on the frequency of past recombination events (Brown, 2002). Through analysis of several sequenced genomes

representing different genes, paralogous sets of genes within a single species and not found in other organisms may be presumed to have derived from a duplication. These paralogous clustered genes, or “lineage-specific expansions”, represent relatively recently duplicated gene clusters since they were generated subsequent to the diversification of the genus analyzed. Analysis of 21 prokaryotic sequenced-genomes for lineage-specific expansions revealed that a significant fraction (5-33%) of predicted proteins in the genomes were resultant of recent gene cluster duplications (Jordan *et al.*, 2001). Thus, duplication of clusters is an important process in evolution.

This analysis also demonstrated that paralogous clustered genes tend to be located more closely on bacterial chromosomes than expected by chance (Jordan *et al.*, 2001). This may reflect a history of tandem gene duplication. Interestingly, since these genes in the paralogous gene clusters did not decay into pseudogenes, these duplicated clusters are believed to reflect genes valuable for adapting to selective pressures. While co-amplification of gene clusters serves as a means for adaptation by increasing protein dosage, this evidence suggests that co-amplification may provide a means for the evolution of the paralogous gene clusters. However, controversy exists concerning whether the evolution of these paralogous genes occurred under neutral or positive selection.

The demonstration that duplications arise spontaneously and independently of selection is consistent with the theory that paralogous genes evolve under neutral selection (Ohno, 1970). According to this neutral evolution theory, once duplicated genes are generated, they are redundant and therefore free to diverge without the loss of genetic information. Random mutations may accumulate on one of the copies and the most common fate of duplicated genes is

the simple silencing, thus forming pseudogenes. However, random mutation may provide an organism with a selectively valuable trait. Such mutations are then retained in the genome.

An alternative hypothesis is that paralogous genes and clusters evolve under positive selection. Although duplications arise independently of selection, tandem duplications and higher amplification of clusters could affect fitness directly, primarily through a protein dosage effect. Support for this positive selection theory has been provided by genomic analysis indicating that lineage-specific expansions arose under purifying selection rather than genetic drift (Kondrashov *et al.*, 2002). If this is the case, the key to making the transition from multiple copies of homogenous genes to functionally different paralogous clusters is that cells must overcome this original selection in order to achieve the long-term advantage of obtaining novel gene functions. Thus, evolution of paralogous genes and clusters could occur in a two-stage evolutionary model. During the first stage, immediately after duplication multiple gene copies would be retained because of the short-term advantage of protein dosage. During the second stage, gene duplication would provide a long-term advantage by allowing the duplication to become fixed in the genome and through divergent evolution enabling the creation of new functions (Kondrashov *et al.*, 2002). Similarly, Henrickson *et al.* (2002) have proposed a model whereby gene amplification is an intermediate step towards the evolution of paralogous genes. In this model, an individual gene or set of genes may be selected for a secondary function, which the gene or cluster is less efficient at performing. Gene amplification may provide a selective advantage by increasing expression of the targeted genes; this allows the organism to perform the secondary function more efficiently. Since the copies are redundant, random mutations may accumulate in some of the copies. By chance, mutations would allow for the gene product(s) to perform the secondary function more efficiently. This would relieve the need for amplification

and the cell would deamplify. During deamplification, selection would require the cell to retain one wild-type copy of the amplicon to perform its primary function and a second mutated copy to perform the secondary function. These theories concerning the involvement of amplification in evolution of paralogous genes may be difficult to prove since genomic analysis is unable to distinguish between a duplication event and an amplification event that is followed by deamplification to two copies.

Whether or not selective amplification is involved in the formation of paralogous gene clusters, it is clear that if a duplicated segment is to be stably maintained there is a strong demand to diverge in sequence identity between the two copies. The high homology and close proximity between redundant copies causes a segmental duplication to be very unstable. Therefore, in order for a duplicated copy to be stably inherited there may be an increased demand to diverge in identity immediately after duplication. Such divergent evolution in the two copies may include multiple nucleotide substitutions, reshuffling of gene order, and other rearrangements such as deletions, insertions, and transpositions. Genomic analysis provides evidence supporting a strong selection for divergence immediately after duplication. Kondroshov *et al.* (2002) found that paralogous genes evolve significantly faster than unduplicated genes with a similar level of divergence. There is also data showing that once segmental duplications are formed occasionally they can promote further rearrangements (Flores *et al.*, 2000).

Conclusion

Spontaneous tandem duplications occur frequently and can be relatively large, encompassing many transcriptional units. Higher amplification of gene clusters may be selected and contribute

to cluster maintenance if the selective conditions are experienced frequently. For tandem duplications to be stably incorporated into the genome, the two segments must diverge significantly in sequence identity. These two processes may be linked together by first providing a short term advantage of increased protein expression during times of stress and then by allowing for the long-term evolution of sets of genes by providing redundant copies for mutation. This linkage may be difficult to prove since it would have occurred over millions of years; however, in the future, the contribution of gene amplification in the evolution of gene clusters will certainly be further investigated. A better understanding of how genome plasticity contributes to evolution would result from a model system for studying gene amplification. The following chapters describe the first steps toward the development of a model system for studying this important process.

CHAPTER 2

GENOME PLASTICITY IN *ACINETOBACTER*:

NEW DEGRADATIVE CAPABILITIES ACQUIRED BY THE SPONTANEOUS

AMPLIFICATION OF LARGE CHROMOSOMAL SEGMENTS¹

¹Reams, A.B. and E.L. Neidle. 2003. *Molecular Microbiology*. 47(5):1291-1304
Reprinted here with permission of publisher, 10/31/03.

Summary

In *Acinetobacter* sp. ADP1, growth on benzoate requires regulation of the *cat* genes by two transcriptional activators. Here, mutants were obtained from a strain lacking both activators by selecting for growth on benzoate medium. The mutants, which arose within three weeks at a frequency of approximately 10^{-8} , carried amplified chromosomal regions (amplicons) encompassing the *cat* genes. Multiple occurrences of low-level expression of *catA* and the *catBCIJFD* operon provided sufficient transcription for growth. The amplicons of four independently isolated mutants varied in size from approximately 30 to 100 kbp of the normally 3.8 Mbp chromosome. Mutants had approximately 10 to 20 copies of an amplicon in adjacent head-to-tail orientations. At the amplicon's chromosomal endpoint, an atypical junction juxtaposed normally distant DNA regions from opposite sides of the *cat* genes. The sequences of these junctions revealed the precise recombination sites underlying amplification. Additionally, amplicon stability was evaluated in the absence of selective pressure. The natural competence of *Acinetobacter* for transformation by linear DNA allowed the development of a powerful new model system for investigating chromosomal rearrangements and for engineering DNA amplifications for wide-ranging applications. The frequent spontaneous amplification of these large chromosomal segments demonstrated the importance of supraoperonic gene clustering in the evolution of catabolic pathways.

Introduction

Gene amplification is an important process in both prokaryotic and eukaryotic organisms that contributes to diverse manifestations including antibiotic resistance, bacterial virulence and certain cancers (reviewed by Romero and Palacios, 1997; Savelyeva and Schwab, 2001; Stark and Wahl, 1984). Here, gene amplification was discovered in the soil bacterium *Acinetobacter* sp. strain ADP1 during the genetic characterization of mutants that spontaneously acquired the ability to grow on benzoate without using two transcriptional activators. These activators, BenM and CatM, are homologous LysR-type regulators that normally enable benzoate to be degraded by increasing the expression of more than a dozen genes (Fig. 2.1) (Collier *et al.*, 1998; Romero-Arroyo *et al.*, 1995).

The complexity of the BenM-CatM regulon underscores the importance of optimizing gene expression for aromatic compound catabolism. Both regulators respond to the catabolite *cis,cis*-muconate (CCM), and BenM also responds to benzoate. Multiple inducers further modulate gene expression, as indicated by the synergistic transcriptional activation of *benA* by BenM in the presence of both benzoate and CCM (Bundy *et al.*, 2002). To test the possibility that BenM and CatM additionally serve to repress the expression of genes for 4-hydroxybenzoate (4HB) degradation during growth on benzoate, mutants were sought that could grow on benzoate without either regulator (Brzostowicz *et al.*, 2002).

BenM and CatM activate the expression of a cluster of chromosomal genes including the *benABCDE* operon, the *catA* gene, and the *catBCIJFD* operon (Fig. 2.1). In a previously characterized *benM* mutant, ACN147, a point mutation strengthens the *benA* promoter and allows high-level *ben* operon expression without requiring a transcriptional activator (Bundy *et al.*, 2002; Collier *et al.*, 1998). As described here, *catM* was deleted from the chromosome of

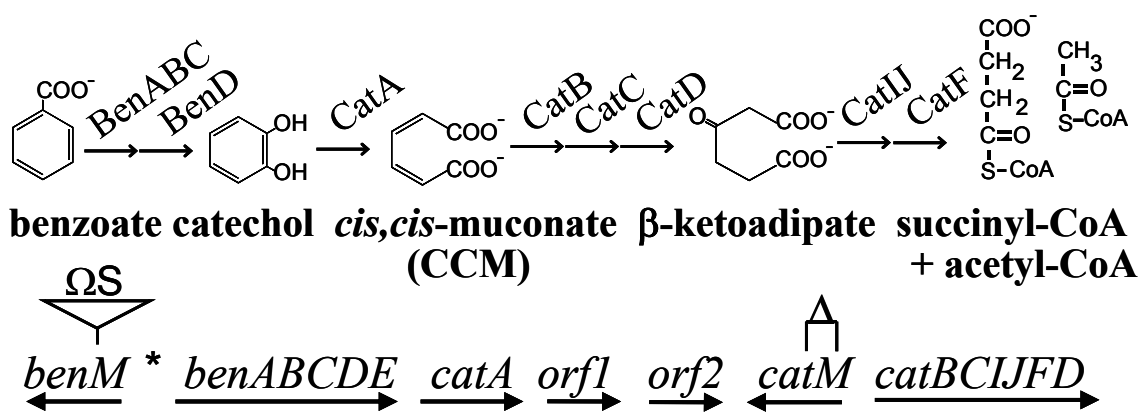


Fig. 2.1 Benzoate degradation by *Acinetobacter* sp. strain ADP1 via the β -ketoadipate pathway. Genetic organization of the chromosome is depicted below the enzymatic reactions catalyzed by the Ben and Cat gene products. The parent strain from which amplification mutants were isolated, strain ACN293, lacks the BenM and CatM transcriptional regulators because of a drug resistance marker inserted in *benM* (ΩS) and an engineered deletion in *catM* (Δ). In this strain, a promoter point mutation (indicated by *) allows the *benABCDE* operon to be expressed without a transcriptional activator (see Table 2.1).

ACN147 to generate a strain lacking CatM and BenM. The resultant strain, which carried the *benA* promoter mutation, should allow high-level *ben* operon transcription but not high-level *cat* gene transcription. It seemed likely that this strain would grow on benzoate only if additional mutation(s) enabled high-level expression of the physically separated *catA* gene and the *catBCIJFD* operon without BenM and CatM. Surprisingly, such mutants were isolated at a relatively high frequency by selection on solid medium with benzoate as the sole carbon source.

As described elsewhere, these mutants were used to study the co-metabolism of benzoate and 4HB (Brzostowicz *et al.*, 2002). In this report, the mutants were characterized genetically. No mutations were identified in the *catA* and *catB* promoter regions. Furthermore, it appeared that no other transcriptional activator replaced the functions of BenM and CatM. Finally, a transformation assay identified DNA fragments in the mutants that could confer the ability to grow on benzoate to the parent strain lacking the transcriptional activators. Analysis of these fragments revealed that gene amplification was responsible for growth on benzoate. In most experimental systems it remains difficult to characterize the endpoints of chromosomal regions that have undergone amplification. As shown in this first report of gene amplification in *Acinetobacter*, a simple transformation assay can determine precise recombination sites involved in spontaneous duplication events that span large regions of the chromosome. Methods for characterizing chromosomal changes are described herein, and the recombination sites in four independently isolated mutants are reported. Moreover, it was possible to use the transformation protocol to engineer a specific amplification mutant.

The soil microbe *Acinetobacter* sp. strain ADP1 should serve as an ideal model system for studying genome plasticity both because of its genetic tractability and its catabolic versatility (Juni, 1978). The significance of the chromosomal clustering of functionally related genes has

been the topic of previous speculation (Young *et al.*, 2001). As shown here, large genomic supraoperonic clusters in *Acinetobacter* readily undergo high-level amplification to confer new catabolic traits. Such amplification and catabolic evolution has important implications for the biodegradation and bioremediation of aromatic compounds that are prevalent in the environment.

Results

Selection of strains that grow on benzoate without BenM or CatM

As a first step in isolating regulatory mutants that grow on benzoate, a parent strain was constructed that lacked CatM, lacked BenM, and expressed the *benABCDE* operon without a transcriptional activator. The latter two features are in ACN147 (Table 2.1), a strain whose *catM* gene became a target for deletion. The wild-type *catM* of ACN147 was replaced with the corresponding allele of pBAC334 (*catM*Δ5293) that has a 446 bp *catM* deletion (see *Experimental procedures*). The resultant strain, designated ACN293, did not grow on benzoate, presumably because of insufficient expression of *catA* and the *catBCIJFD* operon. Southern hybridization methods confirmed the chromosomal *catM* deletion (data not shown).

To isolate mutants with higher *cat*-gene expression, a cellular suspension of ACN293 was spread on solid medium with benzoate as sole carbon source. Within a three-week incubation period, spontaneous mutants were obtained at a frequency of approximately 10^{-8} . This frequency was higher than predicted for separate mutations affecting expression of *catA* and *catB*. Consistent with this inference, the *catA* and *catB* promoter regions of several mutants were amplified by PCR, and all had wild-type sequence. From independent selection experiments, four strains that grew on benzoate (Ben⁺ mutants) were isolated and designated ACN277, 282,

Table 2.1 Relevant strains and plasmids

Strain or plasmid	Relevant characteristic(s) ^a	Reference or source
<i>Acinetobacter</i> strain		
ADP1	wild-type (also designated BD413)	(Juni and Janik, 1969)
ACN147	<i>benM::ΩS4036, benA5147</i> ; point mutation increases <i>benA</i> promoter strength to allow growth on benzoate	(Collier <i>et al.</i> , 1998)
ACN277	<i>benM::ΩS4036, benA5147, ΔcatM5293</i> , amplicon 5277; Ben ⁺ , approximately 30 kbp <i>cat</i> -gene amplicon	This study
ACN282	<i>benM::ΩS4036, benA5147, ΔcatM5293</i> , amplicon 5282; Ben ⁺ , approximately 100 kbp <i>cat</i> -gene amplicon	This study
ACN283	<i>benM::ΩS4036, benA5147, ΔcatM5293</i> , amplicon 5283; Ben ⁺ , approximately 38 kbp <i>cat</i> -gene amplicon	This study
ACN284	<i>benM::ΩS4036, benA5147, ΔcatM5293</i> , amplicon 5284; Ben ⁺ , approximately 50 kbp <i>cat</i> -gene amplicon	This study
ACN293	<i>benM::ΩS4036, benA5147, ΔcatM5293</i> ; Ben ⁻ parent strain from which mutants were selected	This study
ACN382	"X": <i>ΩS5382</i> ; interposon mutagenesis used to map downstream locus involved in generating amplicon 5277	This study
ACN383	<i>benM::ΩS4036, benA5147, ΔcatM5293</i> , amplicon 5383; Ben ⁺ , approximately 26 kbp <i>cat</i> -gene amplicon, engineered amplification	This study
ACN487	<i>recA::Tn5, benM::ΩS4036, benA5147, ΔcatM5293</i> ; replacement of wild-type <i>recA</i> allele of ACN293 with pZR106 DNA	This study
Plasmids		
pUC18 pUC19	Ap ^r , cloning vectors	(Yanisch-Perron <i>et al.</i> , 1985)
pUI1638	Source of cassette conferring SpSm ^r (ΩS)	(Eraso and Kaplan, 1994)
pIGG4	<i>HindIII-EcoRI</i> (11,950-12,892 ^b) <i>catM</i> DNA in pUC19	This study
pIGG6	<i>StuI-SphI</i> (12,381-12,825 ^b) <i>catM</i> fragment deleted from pIGG4	This study
pIGG12	<i>EcoRI</i> (12,892-18,153 ^b) fragment with <i>catC::ΩS4016</i> allele; wild-type <i>catBC</i> DNA from <i>EcoRI-SphI</i> (12,892-14,321 ^b)	(Gaines <i>et al.</i> , 1996)
pZero-2.2	Km ^r , cloning vector; lethal <i>ccd</i> gene for positive selection of recombinant plasmids with cloned inserts	Invitrogen
pZR106	carries <i>Acinetobacter recA::Tn5</i> allele	(Gregg-Jolly and Ornston, 1994)
pBAC200	Plasmid used to isolate the chromosomal <i>catB</i> promoter region of mutants by the gap repair method ^c	(Cospers <i>et al.</i> , 2000)

Table 2.1 (continued)

pBAC201	Wild-type <i>EcoRV-HindIII</i> (7,874-10,764 ^b) fragment ligated to <i>EcoICR</i> and <i>HindIII</i> digested pUC18. Source of <i>Sall-SphI</i> (8,235-8,754 ^b) fragment for <i>catA</i> probe	This study
pBAC334	<i>EcoRI</i> (12,892-18,153 ^b) fragment with <i>catC::ΩS4016</i> from pIGG12 inserted in <i>EcoRI</i> site of pIGG6; the <i>ΔcatM5293</i> allele but not the <i>catC::ΩS4016</i> allele was used to replace chromosomal DNA to make ACN293	This study
pBAC405	19 kbp <i>EcoRI</i> fragment of ACN277 in pZero2.2. "X" region fused to <i>ben</i> (3,760-5,663 ^b) region	This study
pBAC406	5.6 kbp <i>KpnI-EcoRI</i> fragment of pBAC405 subcloned in pZero2.2. "X" region fused to <i>ben</i> (3,760-5,663 ^b) region	This study
pBAC420	1.5 kbp <i>KpnI-XbaI</i> fragment of pBAC405 subcloned in pZero2.2. "X" region fused to <i>ben</i> (3,760-5,663 ^b) region	This study
pBAC426	3.6 kbp <i>NsiI</i> fragment of pBAC406 subcloned in pZero2.2. Contains "X" but not <i>ben</i> DNA	This study
pBAC434	2.2 kbp <i>PstI</i> ΩS fragment of pUI1638 inserted into <i>PstI</i> site of "X" region in pBAC426	This study
pBAC437	4.0 kbp <i>EcoRI</i> fragment with ΩS of ACN382 cloned in pUC19. Isolation of the chromosomal "X":ΩS53824 allele of ACN382 for further characterization	This study
pBAC438	<i>XbaI-NsiI</i> "X" DNA derived from pBAC420 fused immediately upstream of <i>NsiI-SphI cat</i> (7,822-8,754 ^b) DNA in pUC19. Engineered junction used to generate amplification mutant ACN383	This study
pBAC460	3.5 kbp <i>EcoRI</i> fragment of ACN283 in pZero2.2. E-fragment DNA fused to IS1236 ^d immediately upstream of <i>ben</i> (3,715-5,663 ^b) region, see Fig. 2.7	This study
pBAC469	8.7 kbp <i>EcoRI</i> fragment of ACN282 in pZero2.2. D-fragment DNA fused immediately upstream of <i>ben</i> (4,620-5,663 ^b) region, see Fig. 2.7	This study
pBAC509	6.0 kbp <i>EcoRI</i> fragment of ACN284 in pZero2.2. D-fragment DNA fused immediately upstream of <i>ben</i> (3,010-5,663 ^b) region, see Fig. 2.7	This study
pBAC585	1.5 kbp ADP1 <i>NotI-HindIII</i> fragment in cloning vector. Contains E-fragment DNA, immediately adjacent to the <i>NotI</i> site border with the D-fragment, used to make a hybridization probe (JF321)	(Galton et al., 1997)

^aAp^r, ampicillin resistant; SpSm^r, spectinomycin and streptomycin resistant; Km^r kanamycin resistant

^bNumbers correspond to positions in ADP1 DNA sequence in GenBank accession number AF009224

^cMethod of Gregg-Jolly and Ornston (Gregg-Jolly and Ornston, 1990)

^dMember of IS3 family of insertion sequences (Gerischer *et al.*, 1996)

283, and 284. Southern hybridization methods confirmed that each mutant retained the mutated *benM* and *catM* chromosomal alleles of the parent strain (data not shown).

ACN277 DNA responsible for growth on benzoate without BenM and CatM

ACN277 was analyzed to identify mutation(s) in its DNA capable of conferring the Ben⁺ phenotype to the parent strain, ACN293. Plasmid-based libraries of DNA fragments from ACN277 were constructed and maintained in *E. coli* (see *Experimental procedures*). When a ColE1-type vector is used for such libraries, the recombinant plasmids are not well maintained in *Acinetobacter* strains. A fragment from a recombinant plasmid that transforms an *Acinetobacter* recipient usually confers a new phenotype after integration into the recipient's chromosome by homologous recombination (Bundy *et al.*, 1998). This method identifies both *cis*- and *trans*-acting mutations. Since the *Acinetobacter* strains are naturally transformable by linear DNA, individual fragments can be excised from plasmids and used in transformation assays to localize mutated regions (Cosper *et al.*, 2000).

One plasmid, pBAC405, with approximately 19 kbp of ACN277 DNA enabled ACN293 to grow on benzoate. The transforming ability was attributed to a 1.6 kbp fragment on subclone pBAC420. This fragment contained 1 kbp of *benABC* that was identical in sequence to the wild-type region (Fig. 2.2). However, the 5' coding region of *benA* was absent. Instead, a 555 bp region of previously unknown sequence, designated "X", was fused to the middle of *benA* at a site that is recognized by the restriction enzyme *Nsi*I. For DNA from ACN277 to confer the Ben⁺ phenotype, the linear transforming fragments needed to carry the novel junction between the "X" and *ben* sequences as well as some DNA on either side of it.

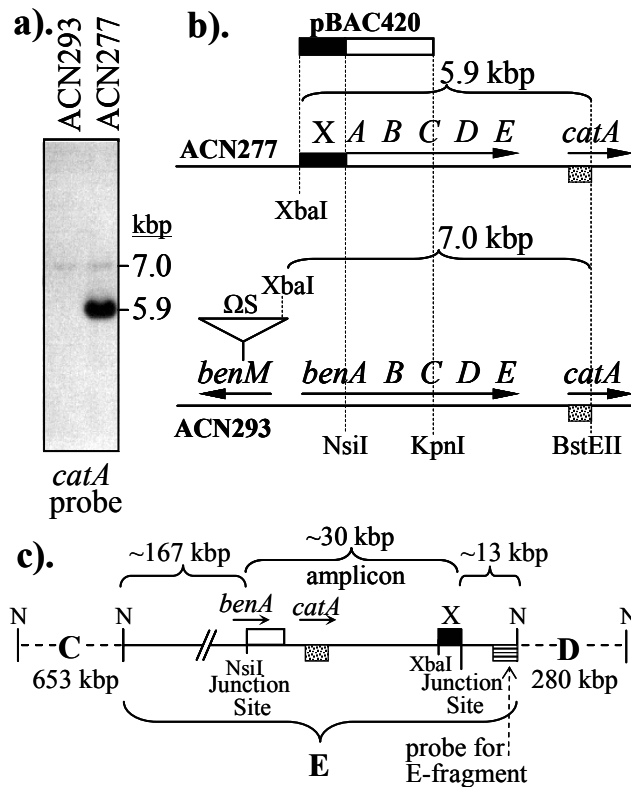


Fig. 2.2 Amplification of a *catA*-containing chromosomal region of ACN277.

a). Southern hybridization of genomic DNA digested by *Xba*I and *Bst*EII with a *catA* probe.

b). Schematic depiction of the *ben*-gene region of plasmid pBAC420 and the corresponding chromosomal regions of ACN277 and ACN293 (not drawn to scale).

c). Schematic depiction of the wild-type chromosome (not drawn to scale). *Not*I restriction sites (N) divide the chromosome into six fragments, three of which are indicated here (C, D, E). The relative positions are shown of the *benA* and sequence "X" sites that recombined as described in the text. Recombination yielded the arrangement in pBAC420 in which a region normally downstream of the *cat* genes (X) is immediately upstream of a portion of *benA*.

Filled black boxes represents the sequence "X" region discussed in the text. Open boxes represents the *ben*-gene region isolated on pBAC420. Hybridization probes for the *catA* gene and the chromosomal *Not*I-generated E fragment are indicated by speckled and striped boxes, respectively. These probes were derived from plasmids pBAC201 and pBAC585 (Table 2.1).

Gene rearrangement and DNA amplification

Hybridization studies showed that ACN277, unlike ACN293 or the wild type, has DNA with the configuration of the pBAC420 insert (Fig. 2.2). For example, chromosomal DNA with sequence "X" fused to the *ben* region should generate a 5.9 kbp fragment when digested with *Xba*I and *Bst*EII. A fragment of this size from ACN277 but not ACN293 hybridized to a *catA* probe. This probe also detected a 7.0 kbp fragment in both ACN277 and ACN293 that is the size of the known *ben*-gene configuration of the parent strain, ACN293. Since ACN277 grows on benzoate, it requires functional *ben* genes. The *ben* region corresponding to the 5.9 kbp hybridization signal should not be sufficient for growth since *benA* is truncated and there is no promoter to express the *benBCDE* genes. In contrast, the *ben* region corresponding to the 7.0 kbp hybridization signal contains an intact *ben* operon including a promoter mutation to allow high-level expression. The relative intensity of the hybridization signals in ACN277 indicated approximately 20 times more copies of the smaller than the larger fragment (see *Experimental procedures* for methods of signal strength analysis). Thus, one chromosomal region in ACN277 was amplified relative to another.

To understand the genetic organization of ACN277, the "X" DNA was studied. A probe of the 555 bp *Xba*I-*Nsi*I fragment of pBAC420 hybridized to chromosomal DNA from ACN277, ACN293, and the wild type (data not shown). PCR primers amplified the wild-type "X" DNA, and its sequence was identical to that on pBAC420. Database homology searches did not reveal its function. The wild-type "X" was mapped to a locus approximately 30 kbp downstream of *benA* and 13 kbp upstream of a chromosomal recognition site for the rare-cutting restriction enzyme *Not*I (*Experimental procedures*). Sites recognized by this enzyme serve as reference points for physical mapping (Gralton *et al.*, 1997). There are six wild-type chromosomal

fragments generated by *NotI* digestion that are designated A to F in decreasing size order. The *ben* and *cat* genes and sequence "X" were all on the *NotI*-generated fragment E (Fig. 2.2c).

Recombination events underlying genomic amplification

To generate the unusual DNA junction in ACN277, DNA upstream of the *cat* genes, in the *ben* region, had recombined with DNA downstream of the *cat* genes, in the "X" region. As illustrated in Fig. 2.3a, further amplification of the resulting duplication could account for the relatively high level of the 5.9 kbp *catA*-containing fragment in ACN277 (Fig. 2.2a). The amplified chromosomal segment, referred to as an amplicon, would correspond to the approximately 30 kbp region between the junction sites in *benA* and region "X" (Fig. 2.2c).

Digestion of the ACN277 genome by *NotI* yielded six chromosomal fragments, the same number as in the wild type and ACN293 (Fig. 2.4). However, strain ACN277 lacked the 210 kbp *NotI*-generated fragment E that normally houses the *ben* and *cat* genes (Gralton *et al.*, 1997). Instead, ACN277 had a novel 810 kbp fragment. Hybridization of this large fragment to a *catA* probe indicated that it corresponded to the E fragment (Fig. 2.4). A second probe, of E-fragment DNA immediately adjacent to the *NotI* site (depicted in Fig. 2.2c), also hybridized to ACN277's 810 kbp fragment (Fig. 2.4). The relative strengths of the *catA* hybridization signals to the normal-sized and enlarged E fragments of ACN293 and ACN277, respectively, indicated that the 810 kbp fragment has multiple copies of *catA*. In contrast, an E fragment region downstream of the amplicon did not appear to be amplified in ACN277 based on the comparative strength of hybridization signals to this probe (Fig. 2.4). Consistent with the estimated 30 kbp size of the amplicon (Fig. 2.2c) and the estimated 20 copies of the *cat* region (Fig. 2.2a), the *NotI*-generated E fragment of ACN277 contained 600 kbp more DNA than that of ACN293.

Fig. 2.3 (next page) Models for gene amplification and the transformation assay.

a). & b). Recombination (dashed X) between a region downstream of the *cat* genes (x y z) and one upstream (a b c) can generate a duplicated region (amplicon). Depending on the position of the recombination event relative to the six wild-type *NotI* recognition sites on the ADP1 chromosome, the amplicons may all reside on the same *NotI*-generated fragment as the *cat* genes as in the top panel (a). Alternatively, as in the middle panel (b), digestion by *NotI* may yield a fragment of the exact size as the amplicon.

c). In the transformation assay, a fragment with the duplication junction from an amplification mutant, such as ACN277, transforms the parent strain, ACN293, to generate the identical duplication of the donor strain. The transforming fragment serves as a template for homologous recombination (dashed X). Although depicted as occurring between sister chromosomes, homologous recombination may instead occur between different regions of the same chromosome.

In each panel, the existence of large-sized perfect homology between the halves of the tandem duplication can result in further amplification by homologous recombination. These models are based on studies of chromosomal rearrangements, as previously reviewed (Roth *et al.*, 1996).

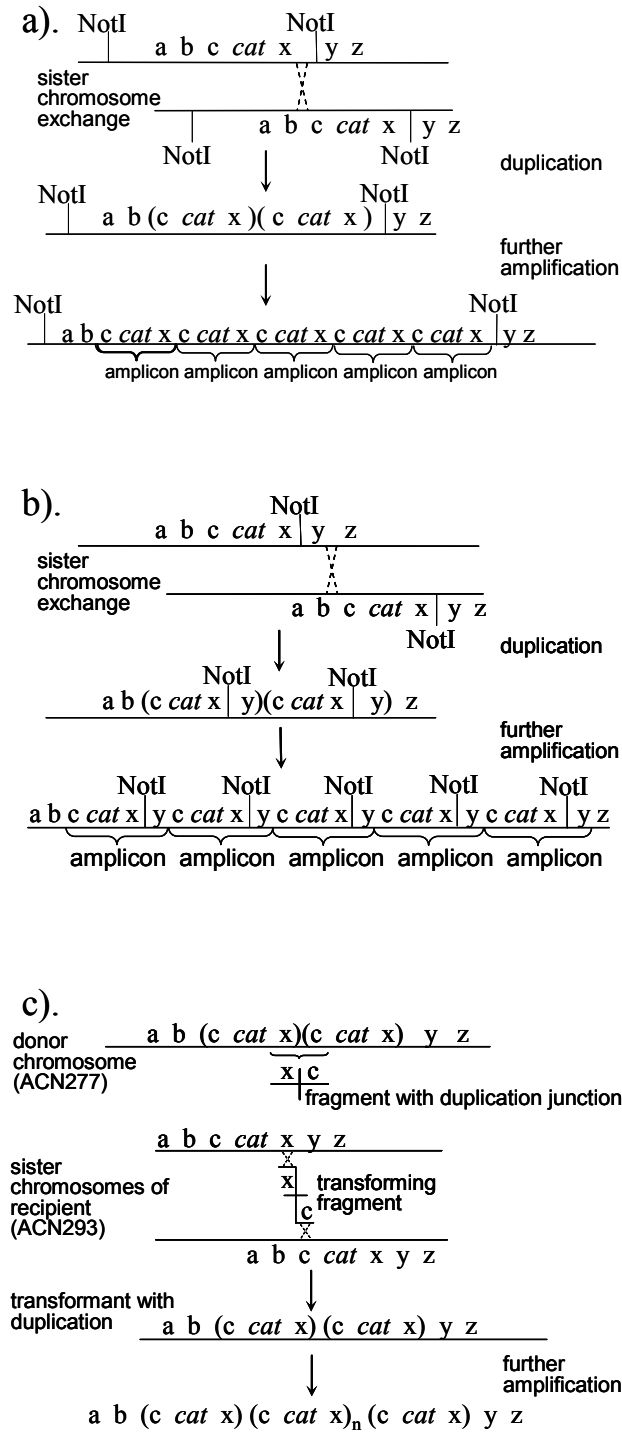


Figure 2.3

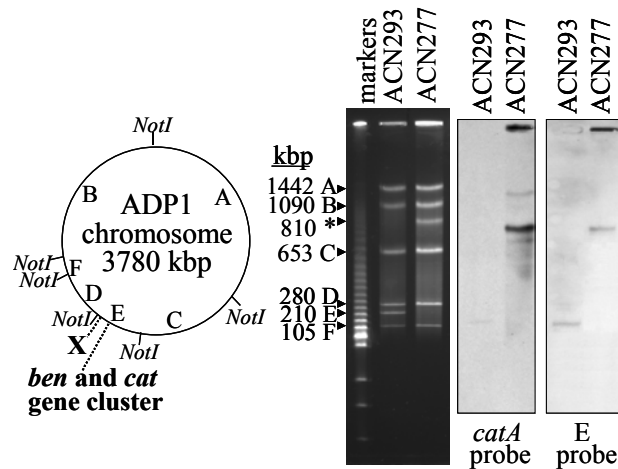


Fig. 2.4 Digestion of the wild-type ADP1 chromosome by *NotI* yields six fragments designated A to F in decreasing size order. These fragments are indicated on a circular map (Gralton *et al.*, 1997) showing the position of the *ben* and *cat* gene cluster relative to the sequence "X" region. The positions and sizes corresponding to these fragments are indicated adjacent to *NotI*-digested chromosomal DNA separated on a pulsed-field gel. An asterisk marks the elevated "E" fragment of strain ACN277 (discussed in the text). This fragment hybridized to *catA* and E-fragment probes (depicted in Fig. 2.2) as shown by Southern hybridization.

The fragment from ACN277 with the novel "X"-*benA* junction represents DNA spanning two adjacent amplicons (Fig. 2.3c). This DNA was initially identified by its ability to transform ACN293 to grow on benzoate. After using this DNA in a transformation assay with ACN293 as the recipient, several of the Ben⁺ transformants were characterized. These transformants were indistinguishable from ACN277 by pulsed field gel electrophoresis and Southern hybridization (data not shown). The transforming DNA generated the same amplicon in the recipient as that of ACN277. It seemed likely that the amplification-inducing effect of the fragment was mediated by homologous recombination (Fig. 2.3c). To test this possibility, the *recA* gene of ACN293 was disrupted to generate strain ACN487 (see *Experimental procedures*). Consistent with the model (Fig. 2.3c), DNA from pBAC420 was unable to transform ACN487 to a Ben⁺ phenotype.

Induction of cat gene encoded enzymes

To understand the contribution of gene amplification to growth on benzoate, the activities of catechol 1,2-dioxygenase (CatA) and *cis,cis*-muconate cycloisomerase (CatB) were measured. These activities were negligible in cell-free extracts of succinate-grown wild-type cultures (Table 2.2). In identically prepared ACN293 extracts, activities were low (Table 2.2). In contrast, these activities were relatively high in succinate-grown ACN277. Multiple copies of weakly expressed *cat* genes should increase gene expression and protein production. ACN277 grew on benzoate at a rate comparable to that of the wild type. Benzoate-grown ACN277 had wild-type levels of CatA and CatB activities (Table 2.2). These activities were slightly higher in benzoate-grown than succinate-grown ACN277 cultures. These differences in enzyme activity might reflect changes in amplicon copy number during different growth conditions.

Table 2.2 Induction of CatA and CatB during growth on succinate or benzoate.

Strain	Carbon Source	Relevant Characteristic(s)	Avg sp activity (nmol/min/mg protein) ^a ± SD	
			CatB	CatA
ADP1	Succinate ^b	Wild type	< 2	< 2
ACN293	Succinate ^b	Parent strain for selection of amplification mutants	4 ± 1	7 ± 1
ACN277	Succinate ^b	Amplification mutant	153 ± 23	283 ± 11
ADP1	Benzoate ^c	Wild type	166 ± 39	372 ± 9
ACN277	Benzoate ^c	Amplification mutant	161 ± 48	365 ± 12

^a The values shown for CatA (catechol 1,2-dioxygenase) and CatB (muconate cycloisomerase) are averages of 3 to 9 repetitions

^b The average generation time was 45 ± 5 min in 3 or more repetitions

^c The average generation time was 90 ± 5 min in 3 or more repetitions

Copy-number stability of the amplicon in ACN277

Most benzoate-grown ACN277 cells had an enlarged 810 kbp E-fragment with approximately 20 amplicons (Fig. 2.4). The amplicon copy number did not change measurably in samples taken at different times during 200 generations of growth on benzoate medium (data not shown). To study amplicon stability in the absence of selective pressure for *cat*-gene expression, benzoate-grown cultures of ACN277 were transferred to succinate medium. Batch cultures were continuously grown and transferred to fresh succinate medium for up to 400 generations. After every 50 generations, cells were prepared for pulsed-field gel analysis of *NotI*-digested genomic DNA. Amplification levels were assessed by Southern hybridization with a probe for E-fragment DNA that lies outside the amplicon (depicted in Fig. 2.2c).

Different fragment sizes in the samples reflected a mixed-population of cells (Fig. 2.5). All the amplicons in a cell should lie on a single *NotI*-generated fragment (Fig. 2.3a). Thus, size variation, which occurred in multiples of approximately 30 kbp units, corresponded to differences in amplicon copy number. Homologous recombination would be expected to delete some amplicons, and there should not be a selection for multiple copies of *cat* genes during growth on succinate. Consistent with these assumptions, the level of amplification decreased over time during growth on succinate (Fig. 2.5). The initial sample contained E-fragments with 12 to more than 20 amplicons each. After 400 generations, most cells contained E-fragments with amplicon copy numbers in the range of one to five.

To determine whether cells with fewer *cat*-gene copies could grow on benzoate, succinate-grown samples were plated on succinate medium after every 50 generations. Individual colonies were then patched to solid benzoate medium. Growth at 37°C was monitored for up to seven days. Although most colonies grew on benzoate (Fig. 2.5b), the time

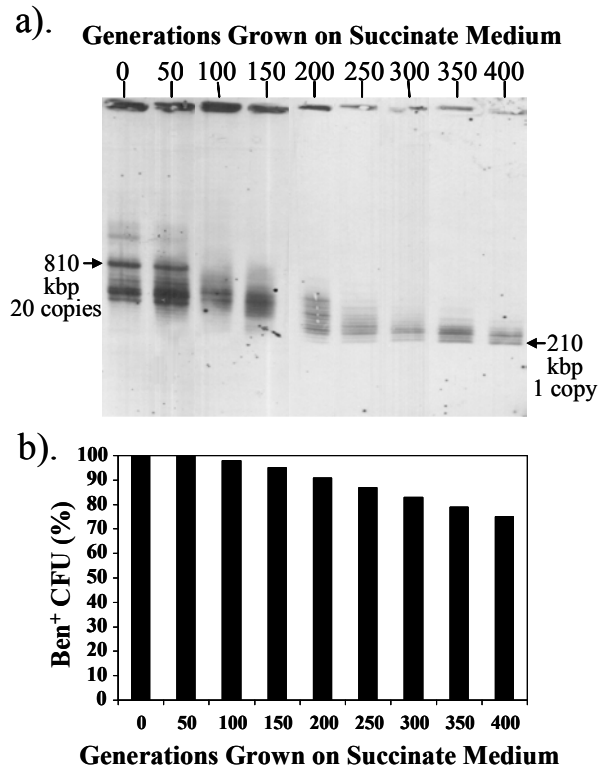


Fig. 2.5 ACN277 culture samples after transfer from benzoate to succinate medium.

a). During growth on succinate medium, samples were taken after every 50 generations and prepared for pulsed-field gel analysis. Genomic DNA, electrophoretically separated after digestion with *NotI*, was used in Southern hybridization studies with a probe to the E-fragment. Arrows indicate fragments corresponding in size to the wild-type (210 kbp) and enlarged (810 kbp) E fragments with approximately 20 copies of the 30-kbp amplicon or a single copy of the *cat*-gene region, respectively.

b). Samples of succinate grown cultures were plated to solid succinate medium. These colonies were patched to medium with benzoate as the sole carbon source to determine the percentage of Ben⁺ CFU (see *Experimental procedures*.)

required for growth varied. There was a higher percentage of slow-growing Ben⁺ patches as the number of generations on succinate medium increased (data not shown). When succinate-grown cells were challenged with benzoate, fewer amplicons per cell corresponded to slower growth rates. Samples of cells from slow-growing patches on benzoate plates were inoculated in benzoate medium for subsequent genome analysis. Pulsed-field gel analysis revealed an increase in the percentage of large-sized *NotI*-generated E fragments in the benzoate-grown population relative to the corresponding succinate-grown population (data not shown). This selection of cells with increased amplicon copy number indicated that multiple copies of the *cat* genes confer selective advantage by allowing faster growth on benzoate.

After growing 400 generations on succinate medium, 25% of the colonies lost the ability to grow on benzoate (Fig. 2.5b). Succinate-grown cells corresponding to the Ben⁻ phenotype were indistinguishable from the parent strain ACN293 when compared by pulsed-field gel and Southern hybridization analysis (data not shown). The complete loss of amplification with the retention of a sole copy of the *cat*-gene region resulted in the inability to grow on benzoate.

Characterization of Amplification Mutants ACN282, ACN283, and ACN284

Three Ben⁺ mutants that were isolated from the parent strain ACN293 independently of ACN277 were also characterized. These mutants grew well on benzoate as the sole carbon source with doubling times of approximately 120 min (ACN282), 100 min (ACN284), and 95 min (ACN283). The comparable wild-type doubling time was 90 min. Growth on benzoate was associated with gene amplification in all the ACN293-derived strains. The majority of benzoate-grown ACN283 cells, like those of ACN277, had enlarged *NotI*-generated E fragments carrying approximately 20 amplicons (data not shown).

In ACN282 and ACN284, the E-fragment was of wild-type size, but there were smaller *NotI*-generated fragments with intense staining patterns suggesting amplification (starred fragments in Fig. 2.6). These fragments, which were approximately 50 kbp in ACN284 and 100 kbp in ACN282, hybridized to a *catA* probe (Fig. 2.6). As illustrated in Fig. 2.3b, *cat*-gene amplification in ACN282 and ACN284 is consistent with the occurrence of a recombination event between DNA upstream of the *cat* genes and DNA downstream of the *cat* genes in a region of the chromosomal D-fragment. Such recombination would incorporate a *NotI* cleavage site in each amplicon, and genomic digestion by *NotI* would yield multiple copies of a fragment identical in size to the amplicon. Thus, the amplicons of ACN282 and 284 were inferred to be 100 and 50 kbp, respectively. The relative intensities of *catA* hybridization signals to the wild-type E fragments and to the amplified fragments indicated that the amplicon copy numbers were approximately 10 and 15 for ACN282 and ACN284, respectively.

Plasmid libraries of *EcoRI*-digested chromosomal DNA from ACN282, 283, and 284 were made in *E. coli*. Plasmids pBAC460, pBAC468, and pBAC509 (Table 2.1) transformed the parent strain ACN293 to grow on benzoate. Linear DNA fragments from each of the mutants could confer a Ben⁺ phenotype to ACN293. Sequence analysis of the fragments revealed novel DNA junctions resulting from recombination between DNA within the *ben*-gene region upstream of the *cat* genes and DNA downstream of the *cat* genes (Fig. 2.7). In ACN283, as in ACN277, the downstream region corresponded to wild-type DNA on the *NotI* chromosomal E fragment. However, in ACN283, unlike ACN277, the insertion sequence IS1236 (Gerischer *et al.*, 1996) was found at the recombination site suggesting its involvement in the chromosomal rearrangement. In ACN282 and ACN284 the downstream regions involved in recombination were on the chromosomal *NotI* fragment D, and there was no evidence of insertion sequences.

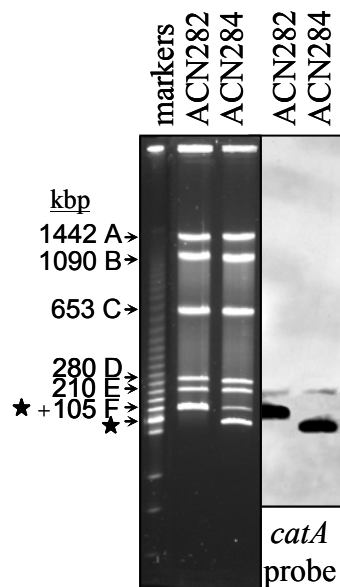


Fig. 2.6 Pulsed field gel (left) and Southern hybridization (right) of *NotI* digested chromosomal DNA of strains ACN282 and ACN284. Stars mark the 100 kbp and 50 kbp amplified fragments of ACN282 and ACN284, respectively. The former fragment was not visually separated from the 105 kbp F fragment under the electrophoretic conditions used here.

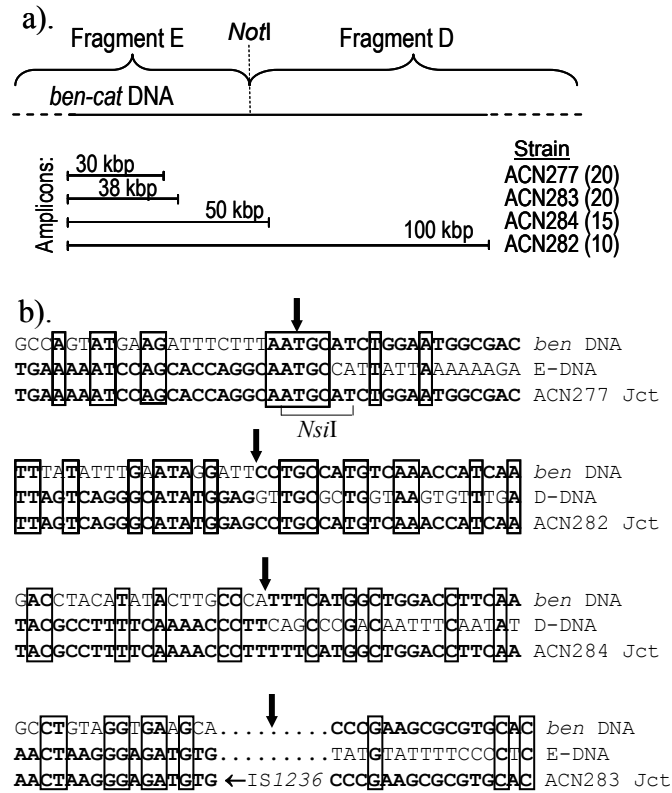


Fig. 2.7 Amplicons and duplication junctions in four independently isolated mutants.

a). Schematic representation of the amplicons in the strains (not drawn to scale). The copy numbers of amplicons in benzoate-grown cultures are indicated parenthetically.

b). Alignments of sequences from the *ben*-genes (top lines), regions downstream of the *cat* genes on either the D or E *NotI*-generated chromosomal fragments (middle lines), and the duplication junctions of mutants (bottom lines). These *ben*-gene regions are all between the transcriptional start sites of *benA* and *catA*. The top lines in the alignments correspond to the following positions in GenBank accession number AF009224: 3740-3779 (ACN277), 2991-3029 (ACN284), 4601-4640 (ACN282), 3699-3730 (ACN283). The middle lines in the alignment correspond to the following positions of the ADP1 genome sequence (www.genoscope.cns.fr): 1,463,987-1,464,026 (ACN277), 1,489,516-1,489,555 (ACN284), 1,522,663-1,522,702 (ACN282), 1,472,835-1,472,866 (ACN283). Arrows indicate recombination sites. Bold residues are identical to those in the duplication junction sequences. Boxes indicate residues that are identical in all three aligned sequences. Jct = Junction.

The precise sequences of the junctions and the approximate chromosomal positions of the amplicons are shown in Fig. 2.7. All four of the junction sequences were unique.

Engineering specific amplification mutants

When used in the transformation assay, the junction sequence DNA appeared to induce amplification in ACN293 by providing a substrate for homologous recombination (Fig. 2.3c). It seemed likely that additional fragments could transform ACN293 to a Ben⁺ phenotype if they had DNA in a similar configuration to that of the junction fragments from the spontaneous Ben⁺ mutants. To test this premise, a region normally downstream of the *cat* genes was fused to one normally upstream of these genes on plasmid pBAC438 (Table 2.1). The "X" DNA of pBAC420 (Fig 2.2a) was fused to a region of *cat* DNA including the *catA* promoter and 465 bp upstream of the coding sequence of *catA*. This plasmid was able to transform ACN293 to grow on benzoate. One Ben⁺ transformant, designated ACN383, was characterized by Southern hybridization and pulsed field gel electrophoresis (data not shown). This strain had an approximately 26 kbp amplicon as predicted by the model depicted in Fig. 2.3c. As observed for ACN277 and ACN283, benzoate-grown ACN383 carried approximately 20 amplicon copies (data not shown). Thus, it was possible to design and generate a specific amplification mutant.

Discussion

Spontaneous chromosomal gene amplification and supraoperonic clustering

Large spontaneously amplified chromosomal segments were detected in *Acinetobacter* at a frequency of approximately 10⁻⁸ during a three-week selection period. The selection conditions demanded that amplicons be greater than 10 kbp in size and that each amplicon contain more

than one transcriptional unit. Despite these constraints, amplification mutants were readily isolated. Furthermore, the bacteria were remarkably flexible in maintaining elevated levels of the amplified regions. In four mutants, the total amplified DNA during growth on benzoate was in the range of 0.6 to 1 Mbp, an amount close to 25% of the normal content of the genome.

Our studies provide direct evidence of the spontaneous and frequent amplification of a chromosomal segment encompassing a large catabolic gene cluster. This evidence supports the premise that gene amplification is a driving force in the evolution of catabolic pathways. The occurrence of large chromosomal groups of genes with related function is common in aromatic-compound degrading bacteria such as *Acinetobacter* sp. strain ADP1 (Harwood and Parales, 1996; Van der Meer *et al.*, 1992). In ADP1, the *ben-cat* genes are adjacent to functionally related *sal* and *are* genes (Jones *et al.*, 1999; Jones *et al.*, 2000). A second supraoperonic cluster lies approximately 300 kbp away and encompasses more than 40 genes associated with the dissimilation of aromatic compounds (Gralton *et al.*, 1997; Parke *et al.*, 2001). Sequence similarities among genes in the two clusters raise questions about the interactions and evolution of the two genetic regions (Young *et al.*, 2001). However, possible connections between gene amplification, evolution, and sequence similarities in these clusters remain to be demonstrated.

The supraoperonic clustering of degradative genes has been associated with the evolution and horizontal transfer of large plasmids and catabolic transposons (Assinder and Williams, 1990; de la Cruz and Davies, 2000; Tan, 1999). These mobile genetic elements contribute to the biodegradation of recalcitrant pollutants that are similar in chemical structure to the naturally abundant aromatic substrates of chromosomally encoded pathways (Van der Meer *et al.*, 1992). Some of these mobile elements are capable of site-specific chromosomal integration and tandem amplification (Klemba *et al.*, 2000; Ravatn *et al.*, 1998). Thus, supraoperonic gene clusters

enable functionally related genes to co-evolve. As shown here, gene amplification plays a significant role in conferring selective advantage for the acquisition of new catabolic traits by allowing the increased production of catabolic proteins through gene dosage effects.

Amplicon endpoints and recombination sites

Four independently isolated amplification mutants were analyzed. The selection conditions required high-level expression of the *catA* gene and the physically separated *catBCIJFD* operon. These genes convert catechol to tricarboxylic acid cycle intermediates via the β -ketoadipate pathway (Fig. 2.1) (Harwood and Parales, 1996). The *benABCDE* genes must also be expressed to permit growth on benzoate. However, the *ben* genes did not have to be amplified in the selected mutants because of a promoter mutation that increases basal transcription of the operon (Bundy *et al.*, 2002; Collier *et al.*, 1998). Since all of the amplicons' upstream endpoints occurred downstream of the transcriptional start site for *benA* (Fig. 2.7), amplification of the *benA* promoter mutation could be detrimental. Such amplification of this promoter could increase the rate of benzoate catabolism relative to that of catechol and thereby cause the accumulation of toxic metabolites. Metabolite imbalances within this pathway can inhibit growth (Gaines *et al.*, 1996).

The selection conditions resulted in one end of each amplicon being within a predictable 5.6 kbp region between the *catA* and *benA* genes. In contrast, the amplicon endpoints distal to the *ben* genes occurred at well-separated positions on the chromosome (Fig. 2.7). These endpoints were all downstream of *catD*. They corresponded to amplicons ranging in size from approximately 30 to 100 kbp. It is not clear what additional constraints or size limitations may

affect the boundaries of amplicons in this system. In recent studies of this system, amplicons as large as 300 kbp were isolated (unpublished data).

The minimal sequence identity between the DNA regions that recombined, shown in Fig. 2.7b, was not characteristic of homologous recombination (as reviewed by Kowalczykowski, 2000). Only in ACN277 were there identical residues in the aligned sequences at the precise recombination site. At this site, aligned identical nucleotides contributed to the recognition sequence for *Nsi*I (Fig. 2.7b). This *Nsi*I site served as a convenient marker for the amplicon endpoint (Fig. 2.2) but had no obvious significance for amplification. Thus, initial sequence comparisons suggest that duplications arose from illegitimate recombination (Ehrlich *et al.*, 1993; Tlsty *et al.*, 1984). Consistent with our results is a two-step amplification process involving first duplication and then RecA-dependent recombination to increase amplicon copy number (Roth *et al.*, 1996, Fig. 2.3a & b). An alternative possibility is that amplification results from a rolling-circle replication mechanism (Petit *et al.*, 1992).

An insertion sequence, IS1236, was identified at the amplicon boundary of ACN283 (Fig. 2.7b). While this element is present in approximately seven copies on the wild-type chromosome, these copies are not usually on the *Not*I-generated E fragment (Gerischer *et al.*, 1996; Gralton *et al.*, 1997, www.genoscope.cns.fr). This element, from another location, appears to have been inserted into E-fragment DNA. The role of IS1236 in the rearrangement and amplification of chromosomal DNA in *Acinetobacter* is not known. However, insertion sequences can provide regions of sequence identity that lead to gene duplication (Roth *et al.*, 1996).

Gene amplification as a substitute for transcriptional regulation

The complexity of the CatM-BenM regulon suggests the importance of optimizing levels of gene expression during growth on aromatic compounds. Usually, CatM and BenM activate the transcription of *ben* and *cat* genes in response to CCM and benzoate (Clark *et al.*, 2002; Collier *et al.*, 1998; Romero-Arroyo *et al.*, 1995). In the absence of these effectors, BenM and CatM repress basal expression of the *benABCDE* operon by blocking the *benA* promoter (Bundy *et al.*, 2002). Similar repression of the basal expression of *catA* and *catB* can be inferred from the sequence similarities and positions of the regulator-DNA binding sites upstream of the BenM- and CatM-activated loci. Here, the absence of CatM and BenM in succinate-grown cultures of ACN293 resulted in slightly higher CatA and CatB enzyme activities than in similarly grown cultures of the wild-type ADP1 (Table 2.2). The increased enzyme activities should reflect de-repressed expression of the corresponding genes.

The CatA and CatB enzyme activities in ACN277 were consistent with the inference that an elevated level of basal *cat*-gene transcription and an increased *cat*-gene copy number contributed to this strain's ability to grow on benzoate (Table 2.2). Growth on benzoate selected mutants that acquired optimal amplicon copy numbers. For example, the majority of the benzoate-grown ACN277 cells each carried approximately 20 amplicon copies (Fig. 2.4). It was possible to draw this conclusion in ACN277, ACN283, and ACN383 (data not shown) because the amplicons were contained entirely on the E-fragment (Fig. 2.3a). Each copy of an enlarged E-fragment came from an individual cell. In contrast, in ACN282 and ACN284 the digestion of one genome by *NotI* liberated multiple single copies an amplicon (Fig. 2.3b). Within the population there were approximately 10 and 15 copies, respectively, of each amplicon relative to

the other chromosomal regions. However, in these latter cases it was not possible to discern the distribution of the amplicon copies in individual cells in the population.

In ACN277 and ACN283, approximately 20 copies of the *cat* genes enabled the mutants to grow on benzoate at a rate comparable to the wild type. In ACN282 and ACN284, larger amplicon sizes correlated with decreased copy number and slightly decreased growth rates. Presumably a balance was achieved between the advantages of multiple *cat*-gene copies and deleterious effects that may result from the amplification of large chromosomal segments. It is not clear in this case whether size *per se* was a key factor. Gene amplification is clearly not the most efficient method for achieving optimal gene expression. Nevertheless, these results demonstrate that the amplification of supraoperonic gene clusters can confer selective advantage under conditions that require novel catabolic activities.

Stability of amplified DNA in the absence of selection for multiple gene copies

After the selective pressure for growth on benzoate was removed, the average number of *cat*-gene amplicons per cell decreased during growth of ACN277 on succinate (Fig. 2.5). This decrease corresponded in general to a slower growth rate of individual colonies when succinate-grown cells were re-exposed to benzoate as the sole carbon source. In comparison to cultures grown for fewer generations on succinate, extended periods of growth on succinate led to individual cells in the population requiring a greater average number of days to form visible patches on benzoate plates. The rate at which the amplification mutants grew on benzoate appeared to depend on amplicon copy number. Continued growth on benzoate enriched the proportion of cells that had higher amplicon copy numbers relative to the starting population. In

contrast, the Ben⁻ cells that were examined after prolonged growth on succinate retained only a single copy of the *cat*-gene region and were indistinguishable from the parent strain ACN293.

Significance of the transformation assay for amplicon analyses

Examples of gene amplification have been documented in many organisms. In prokaryotes, this type of genome plasticity has been especially well studied in *Salmonella enterica* serovar Typhimurium and in *E. coli*, as discussed by Roth *et al.* (1996) and Whoriskey *et al.* (1987). The importance of phenomena associated with gene amplification, such as drug resistance and cancer, has focused attention on the underlying mechanisms, as previously reviewed (Romero and Palacios, 1997; Stark and Wahl, 1984). One important aspect of understanding such mechanisms is to assess the distinguishing characteristics of DNA regions that spontaneously recombine. However, in most systems, it is difficult to locate the boundaries of amplicons that can span any part of the chromosome. Southern hybridization methods help localize the general regions that have been amplified. Nevertheless, hybridization methods are unwieldy for the analysis of large numbers of independently isolated amplification mutants because they do not identify the exact boundaries but only narrow the search for the precise duplication junctions.

In contrast to many traditional methods, the novel transformation assay developed for this *Acinetobacter* system is a simple, easy, and reliable procedure for identifying small DNA fragments whose sequences reveal precise duplication sites. The efficiency of the assay increases by growing the donor strain on benzoate to maximize the copy number of the amplicon within the genome prior to DNA purification. Such growth conditions augment the representation of the desired junction fragment in the plasmid library. The ease with which

chromosomal duplication sites were identified in this study underscores the utility of the assay for the large-scale analysis of amplification mutants.

The RecA-dependence of the assay indicated that homologous recombination generated gene amplification following transformation (Fig. 2.3c). This method of gene duplication appeared to differ from that of the spontaneous duplication events, depicted as the first steps in Fig. 2.3a and 2.3b, which may result from illegitimate recombination. Regardless of mechanism, the ability to generate gene duplication via transformation by a linear DNA fragment allows specific duplications to be engineered for any purpose. As shown by the construction of strain ACN383, linear DNA with the appropriate configuration can be used to amplify regions of the ACN293 chromosome. This method can be used to explore the permissibility of particular chromosomal rearrangements. It can also be used for metabolic engineering in which genes inserted into the ADP1 chromosome, as previously described (Kok *et al.*, 1999), can be deliberately amplified. For example, disruption of either open reading frame downstream of *catA* does not prevent growth on benzoate (unpublished data). Thus, foreign DNA inserted at this site could be co-amplified with the *cat* genes. This *Acinetobacter* system has powerful potential for the analysis of natural amplification processes and for the generation of amplification mutants for a variety of applications.

Experimental Procedures

Bacterial strains, plasmids, and growth conditions

Acinetobacter strains and plasmids are in Table 2.1. *E. coli* DH5 α and Top10 (Invitrogen) were used as plasmid hosts (Grant *et al.*, 1990). Unless otherwise noted, bacteria were cultured at 37°C in Luria-Bertani medium or minimal medium with 10 mM succinate or 2 mM benzoate as

the carbon source (Sambrook *et al.*, 1989; Shanley *et al.*, 1986). Antibiotics were added as needed at the following final concentrations: kanamycin, 40 µg/ml; streptomycin, 13 µg/ml; spectinomycin, 13 µg/ml; and ampicillin, 150 µg/ml. Growth of culture samples was monitored by optical density at 600 nm.

Since it is difficult to distinguish *Acinetobacter* strains that grow very slowly on benzoate as a sole carbon source from those that do not grow at all, the following test was used to classify Ben⁺ strains. Succinate-grown cells were suspended in minimal medium with no carbon source. Different dilutions were spread on benzoate agar medium. Comparable samples were used to count viable cells on succinate medium. Strains were classified as being unable to grow (Ben⁻) if they did not form visible colonies on benzoate plates, which had been inoculated with 10 to 20 cfu per plate, after 7 days of incubation at 25°C. Wild-type colonies form in 1 to 2 days.

Plasmid, library and strain construction

Standard methods were used for DNA manipulations (Sambrook *et al.*, 1989). A plasmid-borne *catM*-deleted allele was made as follows. The *catM* DNA of plasmid pIGG4 (Table 2.1) was digested with *Sall* and *SphI*, treated with mung bean nuclease, and ligated to form pIGG6. This plasmid has wild-type *Acinetobacter* DNA surrounding a 446 bp *catM* deletion. To increase the length of the *catBC* DNA target for homologous recombination in the allelic replacement method, additional DNA was inserted into pIGG6 at the sole *EcoRI* site to form pBAC334. The inserted DNA, excised by *EcoRI*-digestion of pIGG12, was confirmed to have the wild-type orientation in pBAC334.

To make plasmid libraries, *Acinetobacter* chromosomal DNA was purified with an Easy-DNA kit (Invitrogen). Samples of DNA from each mutant were digested with *EcoRI*, ligated to

a similarly digested sample of the pZero-2.2 cloning vector (Invitrogen) and used to transform *E. coli*. This ColE1-type vector, which is lethal in *E. coli* without the insertion of cloned DNA, is designed to select kanamycin-resistant host cells that all carry recombinant plasmids.

Allelic replacement methods for strain construction relied on the natural competence of the *Acinetobacter* strains to be transformed by DNA, including linear DNA. Previously described *Acinetobacter* transformation methods were used (Neidle *et al.*, 1989). To introduce the plasmid-borne *catM*-deleted allele into the chromosome, pBAC334 was linearized by digestion with *Bam*HI and used to transform the recipient, ACN147. Since strains without BenM and CatM do not grow on benzoate as a sole carbon source (Collier *et al.*, 1998), transformants were screened for the loss of the ability to grow on benzoate. After screening approximately 1,400 colonies, one Ben⁻ strain was identified and designated ACN293. Strain ACN293 was later transformed with DNA carrying a transposon-disrupted *recA* allele (Gregg-Jolly and Ornston, 1994). After ACN293 had been transformed with linear DNA (pZR106 cleaved with *Eco*RI), the resultant strain, ACN487, was selected by resistance to kanamycin. The allelic replacement was confirmed by Southern hybridization (data not shown).

Selection of ACN293-derived Ben⁺ mutants and transformation assay to detect DNA that confers growth on benzoate to ACN293

The following method was used to select spontaneous Ben⁺ mutants. After overnight incubation, cells from a 100 ml succinate-grown batch culture of ACN293 were collected by centrifugation and washed three times with minimal medium containing no carbon source to remove residual succinate. Cells were suspended in 1 ml of minimal medium. Serial dilutions of this cell suspension were spread on succinate solid medium to determine viable cell counts. In addition,

100 µl samples of the cell suspension (approximately 10^{10} cells) were spread plated on solid benzoate medium. Plates were incubated at room temperature, approximately 22°C, for several weeks. Incubation at this temperature was designed to prevent desiccation of the plates prior to colony formation. Colonies were streak-purified several times on fresh benzoate medium. This procedure was repeated. Mutants isolated from different batch cultures inoculated from distinct ACN293 colonies were considered to be independently isolated.

To identify mutations responsible for Ben⁺ growth in the absence of BenM and CatM, *Acinetobacter* DNA fragments housed on recombinant plasmids were used to transform the ACN293 recipient as follows. Individual *E. coli* colonies, each presumed to harbor a single recombinant plasmid, were patched first to a "master" plate to allow subsequent retrieval of the plasmid. They were next patched to a succinate medium plate to which a lawn of ACN293 cells had previously been spread. The *E. coli* cells, which do not grow well on succinate medium, can donate plasmid DNA to transform the recipient *Acinetobacter* strain, as shown previously (Bundy et al., 1998). After overnight incubation, the ACN293 cells in the vicinity of the patched donor cells were transferred to minimal medium with benzoate as the sole carbon source. *E. coli* is unable to grow on benzoate.

Plasmids were isolated from *E. coli* cells on master plates that corresponded to regions of Ben⁺ growth. Linear fragments of these plasmids were tested to confirm their ability to transform ACN293. When ACN293 was used as the recipient for transformation by plasmid DNA, a 200-µl sample of an overnight culture was diluted into 5 ml of fresh succinate medium. After incubation for 2 h, DNA (circa 0.1 to 1 µg) was added, and the culture was grown for 3 to 5 h. Residual succinate medium was removed following centrifugation, and cells, suspended in minimal medium with no carbon source, were spread on benzoate agar medium. In positive

transformation assays, colonies were visible within 5 days. The DNA sequences were determined of those regions responsible for conferring Ben⁺ growth.

DNA sequence determination and PCR analysis

Plasmids or PCR-generated fragments were used as templates for DNA sequencing reactions carried out at the University of Georgia Molecular Genetics Instrumentation Facility with an automated sequencer (ABI373A, Applied Biosystems). The *catA* and *catB* promoter regions of mutants were isolated by PCR and gap-repair methods, as previously described (Casper *et al.*, 2000; Gregg-Jolly and Ornston, 1990). The following primer pairs were used to sequence the promoter regions of *catA* and *catB*: 5' AGTAATGATCACATTATGAGC 3' (catA-RN2) with 5' GTATTGAATATTTTAACTTCC 3' (catA-PE1) and 5' GCTGAATTCCCAATTCTT 3' (catMBOP#1) with 5' AAAACTAATGTCTGCGTC 3' (catMBOP#2). For sequencing DNA isolated from the amplification mutants, universal pUC/M13 primers (Promega) were used whenever possible. Additional sequencing and PCR primers were designed as needed and purchased from Genosys Biotechnologies. DNA sequences corresponded to those of the ADP1 genome as indicated in Fig. 2.7. PCR was done with FastStart *Taq* DNA polymerase according to the instructions of the supplier (Roche Applied Science).

Mapping and ADP1 genome sequence

Previously described mapping methods were used to locate the "X" region (Gralton *et al.*, 1997). A cleavage site for the restriction enzyme *NotI* was introduced into a plasmid-borne copy of the target region by inserting a drug-resistance cassette (Eraso and Kaplan, 1994) into the *PstI* site of pBAC426 to generate plasmid pBAC434. Next, the modified "X" DNA was used to replace the

corresponding wild-type chromosomal region of ADP1. After ADP1 had been transformed with linear DNA (pBAC434 cleaved with *Nsi*I), the resultant strain, ACN382, was selected by drug resistance. The position of the engineered *Not*I site within the "X" region was determined relative to the physical map of ADP1 by pulsed-field gel electrophoresis and hybridization. The "X" location, indicated in Fig. 2.2B, was confirmed when genomic sequence for ADP1 became available (www.genoscope.cns.fr). This chromosomal "X" region was isolated from ACN382 by ligating *Eco*RI-digested genomic DNA to a vector and selecting for plasmids that conferred the appropriate drug resistances. The resulting plasmid, pBAC437, was sequenced.

Pulsed field gel electrophoresis and Southern hybridization methods

DNA samples were prepared and electrophoretically separated by the pulsed-field method of transverse alternating field electrophoresis (TAFE, Geneline II, Beckman) as previously described (Gralton *et al.*, 1997). Southern hybridization methods were carried out as indicated by (Gralton *et al.*, 1997). Nonradioactive probes, prepared with digoxigenin and a random-primed labeling system (Genius Systems, Boehringer Mannheim), were hybridized to target sequences, and detected with anti-digoxigenin alkaline phosphatase conjugates and chemiluminescent substrates. Hybridization signal intensities on exposed X-ray films were digitized and calibrated on a UVP Lab Products gel documentation system according to the directions of the LabWorks image acquisition and analysis software.

Amplicon stability studies

Strain ACN277 was transferred from benzoate medium and grown in successive batch cultures for approximately 3 to 20 generations each time with succinate as the carbon source. The initial

benzoate-grown inoculum was washed three times with minimal medium to remove residual benzoate. A volume of 10^{-1} to 10^{-4} of each batch culture was used to inoculate fresh succinate medium until 400 generations had passed. After the cells had grown for successive intervals of 50 generations, samples were used both to determine the viable cell count and to prepare for genomic pulsed-field gel analysis. Colonies from the succinate medium plates, which were used to count viable cells, were patched to benzoate agar medium. The percentage of CFUs retaining the ability to grow on benzoate was determined.

Catechol 1,2-dioxygenase (CatA) and muconate cycloisomerase (CatB) activities

Cell-free extracts were prepared after disruption of *Acinetobacter* cells by sonication as previously described (Shanley *et al.*, 1986). Catechol 1,2-dioxygenase and muconate cycloisomerase activities were assayed spectrophotometrically by the increase or decrease in *cis*, *cis*-muconate concentration, respectively, as indicated by A_{260} (Meagher *et al.*, 1990; Ngai *et al.*, 1990). Protein concentrations were determined by the method of Bradford with bovine serum albumin as the standards (Bradford, 1976).

Acknowledgements

We thank Becky M. Bundy, D. Matthew Eby, and Anna Karls for helpful discussions. We are grateful to Lauren S. Collier for plasmid constructions including pBAC201. This work was supported by National Science Foundation grant MCB 0212604 to E.L.N.

CHAPTER 3

GENE AMPLIFICATION INVOLVES

SITE-SPECIFIC SHORT HOMOLOGY-INDEPENDENT ILLEGITIMATE

RECOMBINATION IN *ACINETOBACTER* SP. STRAIN ADP1¹

¹Reams, A.B. and E.L. Neidle. 2003. Submitted to *Journal of Molecular Biology* 10/29/03.

Summary

A system for studying gene amplification in the bacterium *Acinetobacter* sp. strain ADP1 was used to isolate 105 spontaneous mutants. The method selects for the elevated expression of neighboring transcriptional units in a parent strain lacking two transcriptional activators. Gene amplification can compensate for the activator loss by increasing the copy number of seven weakly expressed genes. Mutant colonies arose from the parent strain at a frequency of 10^{-8} within a three-week period. All but one of these mutants carried tandem head-to-tail repeats of a chromosomal segment (amplicon). These amplicons varied in size from approximately 12 to 290 kbp and ranged in copy number from 3 to more than 30. Gene amplification involved a two-step process in which duplications formed independently of *recA*. Illegitimate recombination fused normally distant chromosomal regions to create novel DNA duplication junctions. These junctions were isolated from amplification mutants using an assay that exploits *Acinetobacter* natural transformability. Sequence analysis of 72 junctions revealed little identity in the recombining regions. Furthermore, multiple independently isolated mutants contained identical junctions. Six different junctions, each found in 2-6 mutants, revealed that some recombination events are site specific. Several recurring junctions were additionally studied using PCR. In each case, the identical duplication present in the mutant was estimated to have occurred in as many as one in a million cells in populations of strains never exposed to selective conditions. These duplications appeared to form spontaneously by a novel type of short-homology independent site-specific process. However, in the absence of *recA*, mutant colonies were not selected from parent cells containing these duplications. Thus, the second gene amplification step most likely depends on homologous recombination to increase amplicon copy number.

These studies support the theory that gene amplification is a driving force in the evolution of functionally related gene clusters.

Introduction

Despite the significant manifestations of gene amplification in hosts ranging from viruses to humans, the underlying mechanisms remain poorly understood. Gene amplification contributes to antibiotic and heavy metal resistance, growth on exotic carbon sources, and increased virulence in microbes (Romero and Palacios, 1997). Additionally, the amplification of mammalian oncogenes contributes to tumorigenesis and metastasis (Stark and Wahl, 1984). This type of genome plasticity enables organisms to adapt to changing environmental conditions by reversibly increasing gene copy number and expression.

DNA amplification may occur by a two-step process that first produces a duplication and then generates additional copies of this duplicated region (amplicon) (Romero and Palacios, 1997; Roth *et al.*, 1996). Although homologous recombination between large direct repeats underlies some tandem duplications, extensive sequence homology is not required for duplication formation. Short direct repeats containing fewer than 15 base pairs have been implicated in duplication formation (Edlund and Normark, 1981; Tlsty *et al.*, 1984; Whoriskey *et al.*, 1987). When recombination occurs between such regions, it is considered illegitimate with weak or no dependence on RecA (Albertini *et al.*, 1982; Yamaguchi *et al.*, 1995). Short-homology dependent recombination can involve base pairing followed by end-joining between two double-strand breaks in sister chromosomes (Roth *et al.*, 1996). Once a duplication occurs, the presence of a long sequence repeat increases the probability that homologous recombination will produce additional copies of the amplicon (Roth *et al.*, 1996).

An improved understanding of gene amplification mechanisms may result from the characterization of DNA sequences involved in the initial duplication step. The key region for sequence analysis is a novel junction produced during gene duplication in which two normally distant genetic regions are fused (fig. 3.1). To date, there have been few large-scale analyses of such junction sites. It is difficult to devise an experimental strategy that allows the isolation of numerous amplification mutants and enables the facile determination of precise amplicon endpoints. In one study using *Escherichia coli*, the duplication junction sequences of thirty independent isolates carrying amplified *lac*-region DNA on an F' episome were determined (Whoriskey *et al.*, 1987). In this region, duplications occurred preferentially by recombination events between DNA segments with short stretches of homology. In a different experimental system involving amplification of the gene 17-18 region of bacteriophage T4, duplications formed between direct repeats of a 24 bp sequence with imperfect homology. These duplications occurred in a sequence-dependent rather than homology-dependent fashion (Wu and Black, 1995; Wu *et al.*, 1995; Wu and Black, 1987; Wu *et al.*, 1991). However, it is unclear whether the conclusions from previous studies may be generalized to gene duplications in diverse organisms and different genetic regions.

As described here, the copy number and size of amplified chromosomal regions were analyzed in 104 independently isolated *Acinetobacter* mutants. Using a recently reported method (Reams and Neidle, 2003), the mutants were selected from a parent strain lacking two transcriptional activators (fig. 3.1). Without these regulators, the parent strain (ACN293) does not express two *cat*-gene transcriptional units at levels high enough to grow on benzoate as the sole carbon source. The spontaneous amplification of a single chromosomal segment encompassing the *catA* gene and the *catBCIJFD* operon confers the ability to grow on benzoate

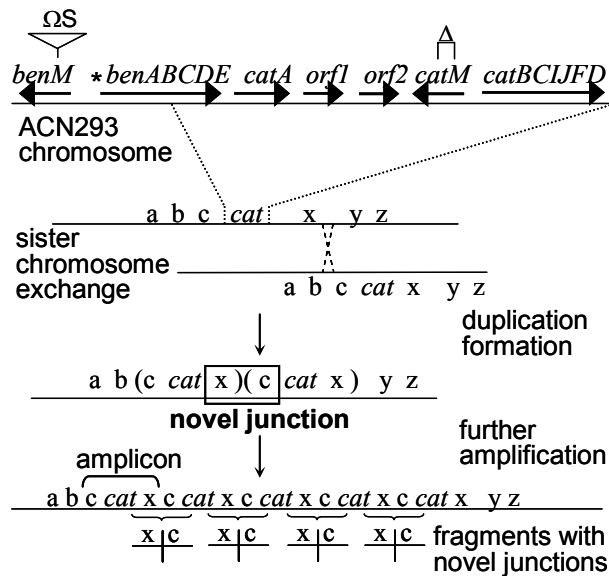


Figure 3.1 Chromosomal duplication generates a characteristic novel junction site juxtaposing normally distant genetic regions. Gene amplification was studied using the clustered *ben* and *cat* genes needed for benzoate catabolism. A parental strain, ACN293, fails to grow on benzoate because of mutations in two genes encoding transcriptional activators, *benM* (disrupted by a drug resistance marker, ΩS) and *catM* (inactivated by a deletion, Δ). A point mutation (marked by an asterisk) allows the *benABCDE* operon to be expressed without a transcriptional activator (Bundy *et al.*, 2002; Collier *et al.*, 1998). Recombination (dashed X) between a region downstream of the *catBCIJFD* operon (x y z) and one upstream of *catA* (a b c) can create a duplication in which the novel junction is shown in a box (x c). The tandem duplication (c cat x) can be further amplified to increase the copy number of this amplicon. A fragment with the identical novel junction is created at each border between amplicons.

(Reams and Neidle, 2003). In each mutant, the upstream amplicon endpoint lies in a region between the promoters of *benA* and *catA*. Presumably, this constraint derives from a point mutation in the *benA* promoter of the parent strain that would cause detrimentally high *ben*-gene expression if present in multiple copies (Bundy *et al.*, 2002; Collier *et al.*, 1998). With the upstream end of the amplicon positioned to include *catA* but not the *benA* promoter mutation, mutants should grow on benzoate if the second boundary lies downstream of *catD* (fig. 3.1).

A transformation assay that takes advantage of the natural competence of these *Acinetobacter* strains enables the determination of the precise endpoints of the amplicons (Reams and Neidle, 2003). This assay involves the construction of chromosomal plasmid libraries in an *E. coli* host for each *Acinetobacter* amplification mutant. These libraries were screened for the ability of recombinant plasmids, which carry the novel junction fragment, to induce gene amplification in the parent strain via homologous recombination (fig. 3.2). This method enables the novel duplication junctions to be identified readily, regardless of the size of the amplicon. Here we report the sequence analysis of the junctions from 72 independent isolates containing tandem gene amplification of the *cat* region. As discussed below, recombination in the *Acinetobacter* mutants occurred at sites in which there was no evident homology. Surprisingly, there were multiple independent occurrences of several identical duplication events despite the lack of any extensive homology present at the recombination site.

Results

Estimation of the cat-region amplicon size and copy number in multiple amplification mutants

Spontaneous mutant colonies that grow on benzoate (Ben⁺) were selected from a parent strain, ACN293, as described in a previous report of four mutants (Reams and Neidle, 2003). This

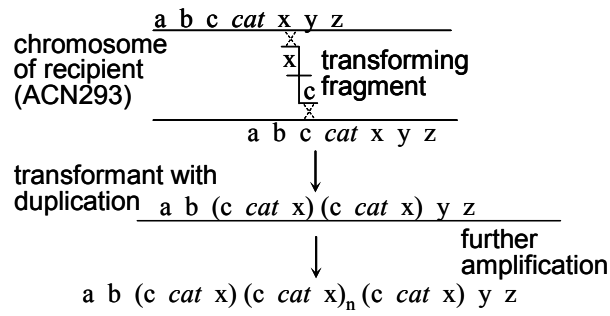


Figure 3.2 Transformation assay detects novel duplication junctions that cause gene amplification in the parent strain. A DNA fragment with the appropriate configuration of regions normally downstream and upstream of the *cat* genes (*x c*) can facilitate homologous recombination in the parent strain (either recombination between different regions of the same chromosome or between sister chromosomes). When such a DNA fragment is used to transform strain ACN293, tandem duplication followed by further amplification results in a transformant that can grow on benzoate due to multiple *cat*-gene copies.

parent strain fails to grow on benzoate because it lacks the transcriptional activators normally used to express high levels of the *cat* genes (Collier *et al.*, 1998; Neidle *et al.*, 1989; Romero-Arroyo *et al.*, 1995). After prolonged incubation on solid benzoate medium, the number of Ben⁺ colonies increased (fig. 3.3). Colonies formed at a frequency of approximately 10⁻⁸ after 3 weeks of incubation. It was likely that these Ben⁺ colonies carried amplified chromosomal *cat*-gene regions (Reams and Neidle, 2003).

To characterize the gene amplification process better, 105 Ben⁺ mutants were independently isolated. Each mutant initially formed a visible colony after incubation periods varying from two to more than thirty days. To assess gene amplification, pulsed-field gel electrophoresis and Southern hybridization methods were used as previously described (Reams and Neidle, 2003). In 104 of the 105 Ben⁺ mutants, there were tandem head-to-tail repeats of a chromosomal amplicon containing the *catA* and *catBCIJFD* genes. Each isolate was determined to carry a single-sized amplicon ranging from approximately 12.5 to 290 kbp. The copy number of the amplicons varied from 3 to more than 30 (summarized in fig 3.4). Some cells carried as much as 2.5 Mbp of additional tandemly repeated DNA on their normally 3.6 Mbp chromosome.

The single Ben⁺ mutant with no evident chromosomal rearrangement arose after two days of incubation, whereas all other mutants initially required at least three days to form a Ben⁺ colony. To identify DNA enabling this mutant (ACN450) to grow on benzoate without *cat*-gene amplification, a chromosomal plasmid library was screened. Analysis of mutant DNA that could transform the parent strain (ACN293) to a Ben⁺ phenotype indicated that the 446 bp *catM* deletion had been partially repaired on the chromosome of ACN450. Although this mutant was not studied further, it appears that the repaired DNA encodes a functional CatM protein that activates *cat*-gene transcription during growth on benzoate.

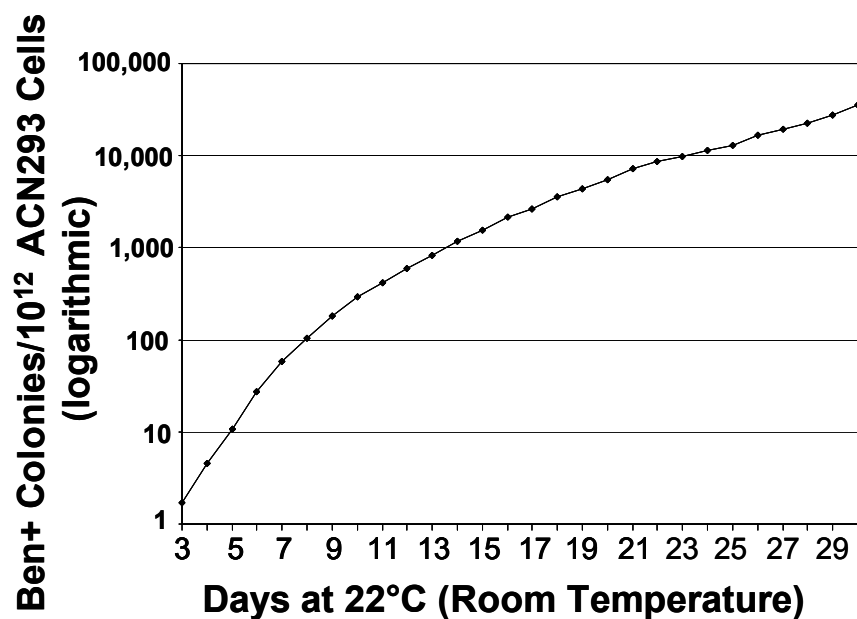


Figure 3.3 Mutant colonies on solid benzoate medium (Ben^+) formed after approximately 5×10^9 ACN293 cells were spread per plate. The total number of visible Ben^+ colonies, counted at the end of each day after inoculation, was compared to the viable cell count determined from a sample of the same culture grown on succinate medium. Each point represents the average value from 20 independently repeated trials with 10 to 20 plates per trial.

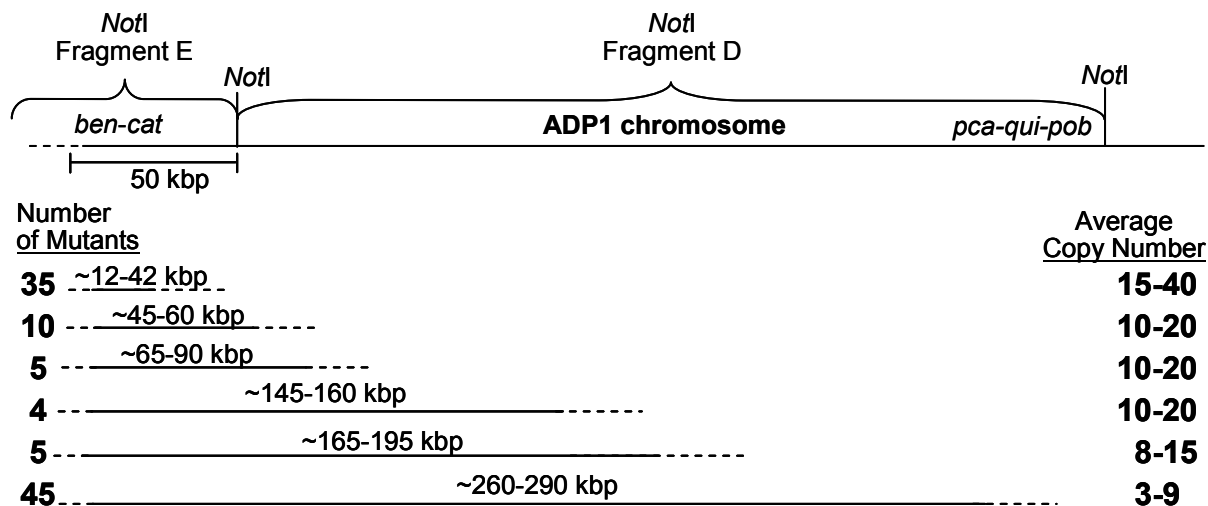


Figure 3.4 Amplicon endpoints relative to the wild-type (ADP1) chromosome. In each case, the upstream boundary of the amplicon was between the promoters of the *benA* and *catA* genes, within a region of the 210-kbp wild-type fragment generated by digestion with *NotI* (designated fragment E) (Gralton *et al.*, 1997). The downstream boundaries were either within this fragment or within the adjacent 280-kbp *NotI* chromosomal fragment D. Dashed lines indicate that precise endpoints are not marked. The total number of mutants with amplicons in a given size range and the average copy number of amplicons for that size range are indicated. When determining amplicon copy number, cultures were grown in benzoate medium, which tends to select for an optimal copy number for the growth condition (Reams and Neidle, 2003).

Sequence analysis of duplication sites in amplification mutants

To identify sequences that recombined during initial duplication formation, a transformation assay (fig. 3.2) was used to isolate DNA carrying novel junctions characteristic of gene amplification (fig. 3.1). With this method, the novel junction sites were obtained from 72 independently isolated spontaneous amplification mutants. The sequences of these 72 junction fragments revealed that in each case the upstream terminus of the amplicon was within a 5531 bp region between the *benA* and *catA* promoters (fig. 3.5). The downstream terminus of each amplicon was in a region located approximately 1.5 to 270 kbp downstream of the *catD* gene. Interestingly, a large portion of the mutants (45/104) amplified an approximately 260-290 kbp chromosomal segment with a downstream terminus in the *pca* operon. The *pca* and *cat* operons are both involved in aromatic compound degradation and share regions of sequence homology (Harwood and Parales, 1996; Young *et al.*, 2001). However, no sequence homology was evident at the exact positions of recombination (discussed below). Furthermore, attempts were made to assess whether homology in the *pca* and *cat* regions could have served as an anchor for distant recombination events. Several different homologous regions were aligned, and the distances were calculated to the precise recombination sites in 27 mutants. This method failed to reveal any significant pattern (data not shown).

The insertional sequence *IS1236* (Gerischer *et al.*, 1996) was found at the recombination site in two different mutants (figs. 3.5 and 3.6), and it appears to have been involved in the chromosomal rearrangement. The upstream and downstream ends of the amplicons were different for each of these two *cat* amplification mutants (fig. 3.6). Furthermore, while multiple copies of *IS1236* reside on the chromosome, neither genome sequencing (http://www.genoscope.cns.fr/externe/English/Projets/Projet_DY/DY.html) nor genome mapping

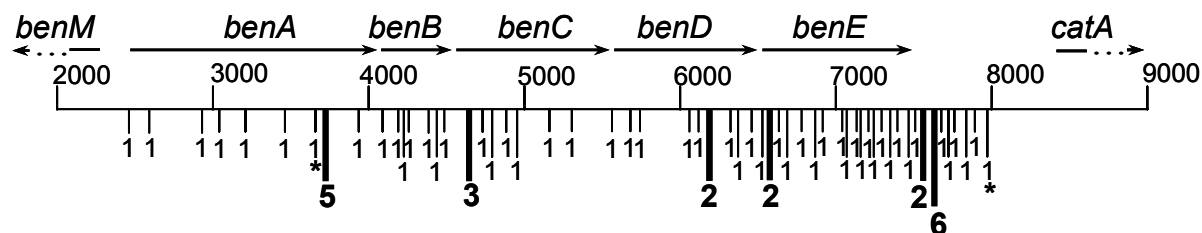


Figure 3.5 Relative positions of the upstream endpoints of amplicons in 72 independently isolated spontaneous mutants. The transcriptional directions and locations of genes are indicated above a scale, marked 2000 to 9000, corresponding to nucleotides in GenBank entry **AF009224**. Where noted by bold lines, the identical junction was found in the indicated number of multiple mutants (from 2 to 6). All other junctions (indicated by a 1) were found in a single mutant. Asterisks mark the duplication sites in two mutants where genetic rearrangements involved the IS/236 element. The amplicon boundary that was farthest upstream (position 2439) occurred immediately downstream of a point mutation in the *benA* promoter of the parent strain, ACN293, such that this mutation is not amplified in the mutant (ACN478, fig. 3.6).

Figure 3.6 (next 3 pages) Alignments of sequences that recombined to form the novel duplication junctions. In each alignment, the middle sequence is that of the junction from the mutant indicated. The bottom line is the parental *ben*-region sequence with the first displayed nucleotide identified by its position in GenBank entry **AF009224**. The top line is the parental downstream region, with the distance (in nucleotides) between amplicon endpoints in parentheses. Identical nucleotides in the aligned sequences are indicated in bold underlined text. Horizontal lines mark the specific recombination sites. For recurring junctions, the number of different independent mutants identified with the same junction is indicated in bold text adjacent to the mutant name. The engineered junction (described in the text) was created in a previous study (Reams and Neidle, 2003) and was never identified in a spontaneous mutant.

Junctions with IS1236: AACTAAGGGAGATGTG ---- TATGTATTTTCCCCTC (38, 365) AACTAAGGGAGATGTG ←IS CCCGAAGCGCGTGCA ACN283 GCCTGTAGGTGAAGCA ---- CCCGAAGCGCGTGCA ben3, 699 AGGTTTGGAAAGCAGATTGA ---- GGCTCAGACAATCAATGCTG (39, 323) AGGTTTGGAAAGCAGATTGA ←IS TTATTTTAAATTTTAGCTTAG ACN377 TATATTTGAAATTTTATTTGTT ---- TTATTTTAAATTTTAGCTTAG ben7, 945	Unique junctions: AGCGCCGTATAATCATAAATGTAGC A GATGCAAAAGCATAAATTTTGCCT (269, 861) AGCGCCGTATAATCATAAATGTAGC A AGCGTGGTAAATAGACCACATATA ACN406 TCGTGGTGACAGCTGCTCGTTCAT A AGCGTGGTAAATAGACCACATATA ben2, 952 ATGGATACGCTTTTCAAAACCTT CAGCCCGACAATTTCAATATGCTCA (87, 666) ATGGATACGCTTTTCAAAACCTT TTTTATGGCTGGACCTTCAATAAAT ACN284 TAATAAGACCACATATCTTGCCCA TTTTATGGCTGGACCTTCAATAAAT ben2, 985
Identical junctions in more than one independently isolated mutant: ATTTGAAAAATCCAGCACCAGGC AATGC CATTATTAAGAAAGAAATCCA (28, 239) ATTTGAAAAATCCAGCACCAGGC AATGC ATCTGGAATGGCGACGCCAGAC ACN277 5 TTCGCCAGTATGAAGATTCTTT AATGC ATCTGGAATGGCGACGCCAGAC ben3, 737 GGCGGATTAGTCAGGGCATATGGAG GTTCCGCTGGTAAAGTGTGACATT (52, 907) GGCGGATTAGTCAGGGCATATGGAG CCTGCCATGTCAAACCATCAAGTAG ACN399 3 TATGAGTTATATTTGAATAGGATT CCTGCCATGTCAAACCATCAAGTAG ben4, 595 AGCCACACACAAACAAAGCTTC A GATGAACTTTGGCTTTTCTCACTC (268, 571) AGCCACACACAAACAAAGCTTC A TTCCGGCTTCTAAAGCGGTGTAAAT ACN400 2 CAGGAGGTATTCACCAATTCAT A TTCCGGCTTGTAAAGCGGTGTAAAT ben6, 103 GATGACGATGCACGGGTGAATCCGA ATGGGGGTGCGATTGCACTGGGCGCA (273, 077) GATGACGATGCACGGGTGAATCCGA TTAGCGGTCTGATTTCACTCTGG ACN445 2 GTTATCGGCCACGATTGCAAGGCTTT TTAGCGGTCTGATTTCACTCTGG ben6, 526 ACCCGGAAACGTTGGTGACCCAGG CATCTTTGTTCAACACGTTGTACCA (270, 135) ACCCGGAAACGTTGGTGACCCAGG TTCGCTTTATTTCTAAGCGCAATTTT ACN408 2 AGACCATTCGATTTATTCATATA TTCGCTTTATTTCTAAGCGCAATTTT ben7, 689 TCATCCAAAGTTTCTGCTGTGTTGCA GCACCACTGGCTTTGGGGGTAACCTC (25, 380) TCATCCAAAGTTTCTGCTGTGTTGCA TGTTCCTATTTCAATTCCTGTTGAAA ACN394 6 AGATGACATATCACTCAGCTTAATT TGTTCCTATTTCAATTCCTGTTGAAA ben7, 749	GATCCGATATTGCTTCAATACTGGCA TT GGTTGAGTCAATTATTGCAAGGT (68, 902) GATCCGATATTGCTTCAATACTGGCA TT TTGGGGGAAACCACCAAAATTAT ACN519 TCCTGATGTACCGTCACTGCAAGAGT TT TTGGGGGAAACCACCAAAATTAT ben3, 168 TGAGTGAATTTAGTCAAGGCTCGAT TAGAAAAGCATTTTTATTTGAGCAA (145, 673) TGAGTGAATTTAGTCAAGGCTCGAT ACTTTCTAAAGCAGCGGAATATAC ACN451 TGGGGTAAATCCGGAAGACCGACCAA ACTTTCTAAAGCAGCGGAATATAC ben3, 490 TGATGGCACCCACCAAGCACTGCGAT T CAACTACACCAAGAACAAATTACGCG (30, 569) TGATGGCACCCACCAAGCACTGCGAT T ACTGGCTCAAAGATATGAAGCAAG ACN489 AAGGCTTGTATTTGGCAGCAGTCAA T ACTGGCTCAAAGATATGAAGCAAG ben3, 947 TTTCACCATGGCGACTCTTCGCCA T GATGCGCGGTGGACATCTGGATAT (274, 135) TTTCACCATGGCGACTCTTCGCCA T TAGACACCATCAGCATCGAACAGA ACN412 GAACGCGATGAATGCTACAGCACTTT T TAGACACCATCAGCATCGAACAGA ben4, 038 GTGATGAGCATTTTGAAGCGTCTGA TATTTTATGAAGCAATCGCGCAT (28, 806) GTGATGAGCATTTTGAAGCGTCTGA GATTTATATCCAGATCGCCAAAGGT ACN482 AAAACCCCAAACTGAGATTTCGCT GATTTATATCCAGATCGCCAAAGGT ben4, 213 CATTTCTTAAGATTTAATTTACTGC TT GACTATTTTAAATGATAAATTAT (31, 666) CATTTCTTAAGATTTAATTTACTGC TT TATATCCAGATCGCCAAAGGTCT ACN518 AACCCCAAACTGAGATTTCGCTGA TT TATATCCAGATCGCCAAAGGTCT ben4, 215 AGCTGGACCTACTCAATAAATAAAA AA TAGCGAAATCGTTATTTTAAAT (36, 646) AGCTGGACCTACTCAATAAATAAAA AA GATCGAGTATTTCGGAATTAAC ACN485 TATATCGAGATCGCCAAGGTCTTG AA GATCGAGTATTTCGGAATTAAC ben4, 242 TTTCACCAAGAGAAGATTTTAAAAA A TTGAGTAGTAGTTAAGTTTGGGT (183, 898) TTTCACCAAGAGAAGATTTTAAAAA A CAGCTATTTTGGCATGTCACGCTA ACN402 CTCAGTTTCCGCTATAAAACAGTT A CAGCTATTTTGGCATGTCACGCTA ben4, 404
Unique junctions: TAGCGTCAAAACTGCCCTAAAGAG GTCGTGCAAAATAATTTAAATGAC (41, 316) TAGCGTCAAAACTGCCCTAAAGAG TATAAACAAGCAAGAAAGACAAGA ACN478 GAACCTTATTAAACATTCTTTTGA TATAAACAAGCAAGAAAGACAAGA ben2, 413 TCATCTAATGGTAAAGAGTTGTGG AA AGCACGTCTTCTGTGGGTAGTCA (288, 770) TCATCTAATGGTAAAGAGTTGTGG AA GGATATGCCAGTATTCCTGTCAT ACN521 TTTCTTATTTGGAAATGTTGGAGA AA GGATATGCCAGTATTCCTGTCAT ben2, 632	AACCTAAATTTTGTATTTAGTATG A ATATTTCTTTTATGAAAGTTGAA (147, 439) AACCTAAATTTTGTATTTAGTATG A TGTAATCGATTCTCAGGTGAACA ACN510 TACAGCTATTTTGGCATGTCACGCT A TGTAATCGATTCTCAGGTGAACA ben4, 428

Unique junctions:

ACGATTTAGTTAAAGGGCTTTAGC | TGGACAAGGAAGAGTCGCAATGA (66, 727)
 ACGATTTAGTTAAAGGGCTTTAGC | AACAAACCAAAATCTTGAGCAAGT ACN449
 CGCTATGTAATCGATTTCTCAGGTG | AACAAACCAAAATCTTGAGCAAGT ben4, 449

GAACCTTGAATCGTGCCGCAAGGCAAT | CTGGCTTGTCGTATACAGGCCGCC (272, 820)
 GAACCTTGAATCGTGCCGCAAGGCAAT | ATACCGTCAGCAATCAATATTCC ACN512
 TCAGGTGAAACCTTATCGATGCAAGC | ATACCGTCAGCAATCAATATTCC ben4, 688

GCAATTTATACAAAACGGCAACCAA | ACTTCGATTAAATTTTCAACAATG (270, 889)
 GCAATTTATACAAAACGGCAACCAA | TATGACATGCCTGAAGACAATTACA ACN515
 CCGTGCTTTTGTGAATCGGGCAAC | TATGACATGCCTGAAGACAATTACA ben4, 772

CAATCCGAAGCTTAAAGCCATGTTA | TGG | TGTGGATACCATGCCGCAAAAC (274, 187)
 CAATCCGAAGCTTAAAGCCATGTTA | TGG | CATGTCAATGCCGTCCAACTTC ACN517
 GAAGAAGCGCAGCAGGCTACGTTT | TGG | CATGTCAATGCCGTCCAACTTC ben4, 842

CAAGCAGATTTGGTGTTTATCAGT | I | AAATGATCTACGCTATATGGGGTT (30, 216)
 CAAGCAGATTTGGTGTTTATCAGT | I | TCAGGCGTCTTCTGAGGTATGTAA ACN411
 CCAACTTCAGATGCTGTATTTCAAAA | I | TCAGGCGTCTTCTGAGGTATGTAA ben4, 884

ACGAAGCGGTTGCACAATACCAGC | I | CAAAGCACGTGCCAAGATCATTGCG (274, 304)
 ACGAAGCGGTTGCACAATACCAGC | I | AGCTCACAACCAGGCAATCGCTTAA ACN524
 CACGGAACACGCTCGTATTGCTT | I | AGCTCACAACCAGGCAATCGCTTAA ben5, 081

CATGTCAACAAAAAGGTGAGCCGA | A | GATTGTAGCTGAACGACGTATCC (273, 164)
 CATGTCAACAAAAAGGTGAGCCGA | A | CCAATTTGGTAGTTTTATATCTGCGT ACN523
 CAGGCGACAAAAAGGTGAGCTTACTGG | A | CCAATTTGGTAGTTTTATATCTGCGT ben5, 197

ITGGAGCACAGGATTTAATACCACT | I | TGCTGAGTTCITATCGTTGGCGTA (37, 598)
 ITGGAGCACAGGATTTAATACCACT | I | GCGGACCGGTTCCTATGGTGAAG ACN490
 ITGGCGGTGAAGTTGATGTGTATCTG | I | GCGGACCGGTTCCTATGGTGAAG ben5, 519

AATGATCAGACAAGGGAATTTCTACTG | AA | TGACAAGGAAATAGTATTTCA (271, 464)
 AATGATCAGACAAGGGAATTTCTACTG | AA | AATTCTCTGCCAACTAATCTCA ACN407
 ATTCAACCGGCGAATTTTATTTGA | AA | AATTCTCTGCCAACTAATCTCA ben5, 598

ACAGGCACAGGTTGGGGTAATTCCT | C | AATCTGCTGCAACTGTGATCGAAC (274, 259)
 ACAGGCACAGGTTGGGGTAATTCCT | C | GTTTTGAACATAAAGTTGTGATTG ACN446
 GGAGACCGATAIGAATTCGACACAA | C | GTTTTGAACATAAAGTTGTGATTG ben5, 651

CGACCAGAATAATTTTGGATTTCAT | ATTGCGCTACAAACATACATGCTTT (177, 617)
 CGACCAGAATAATTTTGGATTTCAT | CAGAAATGCTAAAACATCAACAGGG ACN511
 CTCTGGTGCTGCGGTGCAATCTTC | CAGAAATGCTAAAACATCAACAGGG ben6, 024

Unique junctions:

GCCATTTTCATCGCAATACCATCTAT | ATGCTTTTACAGGAAGCGGGCTGTC (107, 682)
 GCCATTTTCATCGCAATACCATCTAT | AGGGCACTATGTAAATGTGTCTCA ACN418
 TTCCAGAAATGCTAAACATCAACA | GGGCACTATTGTAAATGTGTCTCA ben6, 046

TTCAAATGAATATGATTATGGTG | GACCAAA | ACAAAATGTGCAACACAGAAC (158, 698)
 TTCAAATGAATATGATTATGGTG | GACCAAA | CCATTGATCGAAGCTTTTTA ACN443
 AAGTCTGGATGCAACAAGTGGT | GACCAAA | CCATTGATCGAAGCTTTTTA ben6, 283

CTTTACTCGCAGAAGGTAAACCTAC | CCTGAACCTTIGATGGCAAAAGACTAC (271, 261)
 CTTTACTCGCAGAAGGTAAACCTAC | GCGATTACCTTTTATAGCCTCAGATG ACN522
 TGGCAGTATTGATGAACAAGTCAAT | GCGATTACCTTTTATAGCCTCAGATG ben6, 341

GATTTTACAGTCTAGATGATCTGTC | CATT | TTTTCAATAAACCTGACTTAT (45, 582)
 GATTTTACAGTCTAGATGATCTGTC | CATT | ATTTAAAGGAGACTGCAAGG ACN457
 AGGCGATCAGGGATGATCTGATGGT | CATT | ATTTAAAGGAGACTGCAAGG ben6, 431

AGCAATGGTAATCGCTTCATAACCA | I | CCAGTACCATGACTGAATAGATCA (269, 996)
 AGCAATGGTAATCGCTTCATAACCA | I | TTAGCGGTCTGATTTCACTACTCG ACN444
 CGTTATCGGCCACGATTCAGGGCTT | I | TTAGCGGTCTGATTTCACTACTCG ben6, 525

GACCGTATGACGCAATTTTTGTTGA | AACGGTTTATCGATCATTTACGTA (192, 528)
 GACCGTATGACGCAATTTTTGTTGA | TCGGTTTAAACACCTGTAGTGACC ACN480
 CCGTGACAGGTATTTTATATCTAT | TCGGTTTAAACACCTGTAGTGACC ben6, 669

AACAGTCGGGCGGTAAGTATGCGCT | GTGTTTCAATGTGTATCGGCCTGGGT (272, 938)
 AACAGTCGGGCGGTAAGTATGCGCT | TTATATTACCGCGGCAATTGTGATT ACN453
 TTAGTTTAAATGAAGCCGTGCGCG | TTATATTACCGCGGCAATTGTGATT ben6, 768

TGAATCCGCACCTGTTTCTATCCGA | A | TCTTTATTTGGTGTATTCAAGTGGCG (189, 226)
 TGAATCCGCACCTGTTTCTATCCGA | A | GGGAAATGCTGCTGGCATGATGGC ACN514
 GACAAACTACTCAAAATGGATACCGC | A | GGGAAATGCTGCTGGCATGATGGC ben6, 845

AGTATGGTGCAAGAGCATGAGCGTA | GTTTAGGCGCTTGGCATGCAGAATG (273, 798)
 AGTATGGTGCAAGAGCATGAGCGTA | TGAATCGTGTATTAGCATGTTGCTGTG ACN447
 TTCTTCCACCAGTACTTTACCGC | TGAATCGTGTATTAGCATGTTGCTGTG ben6, 926

AGTAGCTTATATCGAGCGCCGTTTG | AAAGTTCAACTCGGTCTAATCATGC (263, 618)
 AGTAGCTTATATCGAGCGCCGTTTG | TCTGATCAGTCGACGCTTTTACCA ACN455
 TCGTGTATTAGCATGTTGCTGTGTTA | TCTGATCAGTCGACGCTTTTACCA ben6, 954

AGATCAATGGCCCTGAAATGCGCC | TGCAATCGTATTTTCAAACTCATTA (274, 227)
 AGATCAATGGCCCTGAAATGCGCC | ACTGCATGCTTTGGGTACTTATTTG ACN399
 ATCAGTCGACGCTTTTCAAGGAGCTT | ACTGCATGCTTTGGGTACTTATTTG ben6, 983

Unique junctions:

CGCGCGTGAACGAAGACTAAACCACA | TTTAAGTGAACCCAAAGCATGGT (18, 867)
 CGCGCGTGAACGAAGACTAAACCACA | GTAGCTTAACAGGGCAGTTTCTGC ACN415
 AAATTTGGCGCTGCCGCTAAATTTTG | GTAGCTTAACAGGGCAGTTTCTGC ben7, 141

GCACTTGCGGGTAACCGTTTCAAT | AATGAGTCCAICTTGTTCAGCTCG (12, 274)
 GCACTTGCGGGTAACCGTTTCAAT | TATGGCCATTITAAACTCAAGTGGG ACN417
 GCTTAACAGGGCAGTTTCTGCCTGG | TATGGCCATTITAAACTCAAGTGGG ben7, 170

TCGGCGTGGGTCAGGGGATTGCCT | CATTATTGAGCGGGTTTGAGGGGCG (272, 533)
 TCGGCGTGGGTCAGGGGATTGCCT | ACCAATTATTGCCGCCGCAAGCTTG ACN486
 TCAGTGGGTATTAACACACCCGCCAA | ACCAATTATTGCCGCCGCAAGCTTG ben7, 212

TAGTAATAATATTGCGACCTGCGTT | AATGAGTTCACCTGAAGCATAATAA (265, 661)
 TAGTAATAATATTGCGACCTGCGTT | CGCAAGCTTGGCTCTGTTGTTGCT ACN413
 CACCCGCCAACCAATTATTGCGCG | CGCAAGCTTGGCTCTGTTGTTGCT ben7, 227

ATTCCGAGTAATGCATGTTTTCTTT | CATATCTAATTTAGCCCCAAATTA (16, 086)
 ATTCCGAGTAATGCATGTTTTCTTTA | AGCTTGGCCCTCGTTGTTGCTGCC ACN479
 CCGCCAAACCAATTATTGCGCGCGCA | AGCTTGGCCCTCGTTGTTGCTGCC ben7, 230

GCGACGATTGAGGATGTGATTCTT | GG | GGAACACACGGCGGTGCACAAGT (27, 749)
 GCGACGATTGAGGATGTGATTCTT | GG | TGGATTACCATTTGACTGGCC | ACN483
 GGCCTCGTTGTTTGTGCCTTTGCA | GG | TGGATTACCATTTGACTGGCC | ben7, 261

AAAGAGATGTGACTGAAATCTTCAG | I | GATCGCGCTTTAGTGTTCATATATG (272, 530)
 AAAGAGATGTGACTGAAATCTTCAG | I | ATGTATGGGTAAAGATGCACATGA ACN507
 GTACTGGCCCTCATTAACAGCCACT | I | ATGTATGGGTAAAGATGCACATGA ben7, 301

GAAATTACAGCATCATGATTACCA | AT | CGTCAAGGCAAGACATTATCTGT (273, 815)
 GAAATTACAGCATCATGATTACCA | AT | ATTGCTGGAAATAGCAATGGTTI ACN488
 GCACATGAACCTCAAGACAACGCT | AT | ATTGCTGGAAATAGCAATGGTTI ben7, 343

Unique junctions:

TTTTAGATACTGCATGTATTTTCAAT | GTCGTGATGCGCTGGCCATTGTACA (272, 767)
 TTTTAGATACTGCATGTATTTTCAAT | TGGTTTGTTTTATGTTCTAGGAGGT ACN404
 CAAACGCTATATTGCTGGAATAGCAAA | TGGTTTGTTTTATGTTTTAGGAGGT ben7, 360

CAATGACATTTCCATCTCGAAGTAG | A | ACAGGTGGGTTTTCIACAATIGCT (57, 696)
 CAATGACATTTCCATCTCGAAGTAG | A | GTATTGTAGCGTTTTCAGCTTAT ACN477
 TGTITTAGGAGGTGTGTTGCAAGC | A | GTATTGTAGCGTTTTCAGCTTAT ben7, 399

CTTGATATTTCGTTTGGATATTGCTT | G | ATGGCTTGTCTGCAAGAAATTTGA (16, 557)
 CTTGATATTTCGTTTGGATATTGCTT | G | TTTGAGTAGITCAATGCATGACAA ACN456
 TTTATCTCTGTGAAACCACGATAG | G | TTTGAGTAGITCAATGCATGACAA ben7, 783

GGGGCTGGGGCCGATCTACACCCCA | AC | CGGTTTTGGCACTTTACTCGCAG (269, 760)
 GGGGGCTGGGGCCGATCTACACCCCA | AC | AATTTTTTAACCTCTGAATATTG ACN484
 ATAGGTTTGTAGTAGITCAATGATG | AC | AATTTTTTAACCTCTGAATATTG ben7, 804

GACTACGATAATGTGAAGATACAAT | AAAGTAGCGAATGACCTCAGGATGG (39, 697)
 GACTACGATAATGTGAAGATACAAT | CTCTGAATATTCCTTTTATGAAGG ACN513
 AGTTCAATGCATGACAAATTTTTTA | CTCTGAATATTCCTTTTATGAAGG ben7, 816

TTGTAACCACCGATCATGTCAACCA | AAAAGGTGAGCCGAAGATTGTAGCT (270, 515)
 TTGTAACCACCGATCATGTCAACCA | TTATGAAGGTATTTAGCGATATCTG ACN405
 ATTTTTTAACCTCTGAATATTCTT | TTATGAAGGTATTTAGCGATATCTG ben7, 832

ITCATCAATCCGAAGCTTAAAGCCA | I | GTATGGTGTGGATACCATGCCGCA (271, 147)
 ITCATCAATCCGAAGCTTAAAGCCA | I | ATGAGCTAAATTTACTTTTTTAAA ACN520
 IGCTCGACCATAGTAATGATCAGAT | I | ATGAGCTAAATTTACTTTTTTAAA ben7, 880

Engineered junction:

ATTTGAAAAATCCAGCACCAGG | CAATGC | CATTATTAATAAAGAAAAATCCA (22, 177)
 ATTTGAAAAATCCAGCACCAGG | CAATGC | ATGACAATTTTTTAACCTCTGA ACN383
 ACCACGATAGGTTTGAGTAGTT | CAATGC | ATGACAATTTTTTAACCTCTGA ben7, 798

of the wild-type strain identified a copy of this element in the vicinity of the regions that recombined in either mutant (Gralton *et al.*, 1997).

In the remaining 70 junctions, recombination occurred between regions upstream and downstream of the *cat* genes, and no other rearrangements were noted. By aligning each novel junction sequence with the corresponding wild-type upstream and downstream regions, the sequence identity at the precise recombination site was found to be minimal (indicated in fig. 3.6 by bold underlined text). The number of identical nucleotides at the precise recombination site ranged from 0 to 7 bases (marked in bold underlined text between horizontal lines in fig. 3.6). In more than half of the mutants there were no identical bases at the recombination site (36/70). Many mutants contained only a single identical base (18/70) or two identical bases (8/70) at the recombination site. Three individual mutants had three, four or seven identical residues at this position. One junction contained five identical residues at the recombination site, and this exact junction was identified in five independent isolates. In some mutants there were short regions of broken homology near the recombination site. However, the sequence patterns in all cases were indicative of illegitimate rather than homologous recombination (Ehrlich, 1989).

Despite the absence of evident homology, multiple independently isolated mutants contained identical junctions (figs 3.5 and 3.6). Six different recombination events occurred in precisely the same positions in more than one mutant. The most common recurring junction, identified in six independent isolates, completely lacked sequence identity at the exact site of recombination. This independent recurrence of identical junctions is indicative of sequence specificity in the recombination events.

Specific tandem duplications in unselected populations

PCR methods were used to study the presence of specific junctions in various cell populations grown under nonselective conditions. Primers targeting the recurring junctions amplified template DNA containing the corresponding tandem duplications. This method tested the spontaneous occurrence of specific tandem duplications in populations of the wild-type strain (ADP1) and the parent strain (ACN293) that were grown on succinate and never exposed to benzoate.

Primer sets were individually tested for the ability to amplify four of the six recurring junctions. Results from one such experiment are shown in fig. 3.7. PCR products of the expected size for the targeted junction were observed, which indicates the presence of the corresponding tandem duplication in the population tested. To estimate the relative abundance of this specific duplication in different unselected populations, the concentration of the template DNA was standardized and subjected to a series of 10-fold serial dilutions. Samples spanning a wide range of DNA concentrations were used as template for separate PCR reactions. As shown in fig. 3.7d, template DNA from the amplification mutant ACN277, which contains the primer-targeted junction, yielded a PCR product when the template DNA had been diluted 10^7 fold (to a final concentration of 3 fg in the reaction). In contrast, PCR products were identified only in more concentrated samples of DNA from the parent (ACN293) and wild-type (ADP1) strains. In these latter two strains, a PCR product was identified only when the template DNA concentration was 3 ng or higher.

To determine the DNA sequence of these approximately 950-bp fragments, PCR products were cloned from reactions corresponding to the three different strains. In all cases, sequence analysis confirmed that the primers amplified DNA fragments with the identical sequence to the

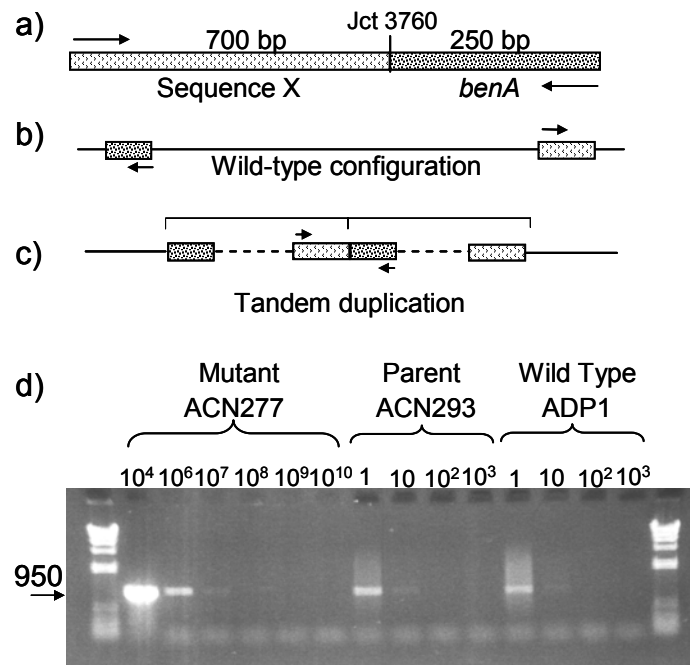


Figure 3.7 Identification of a specific tandem duplication in mutant and wild-type populations. The appropriate PCR primers (arrows) should amplify a 950 bp junction fragment of strain ACN277 (a). The junction is identified by the position (3760) in GenBank entry [AF009224](#) of the upstream endpoint of the *ben*-gene region. This *ben* region in the junction lies immediately downstream of a region (sequence X) that is normally distant. The primers should not amplify the wild-type chromosomal arrangement (b), but they should yield a PCR product in strains carrying this specific tandem duplication (c). PCR reactions were carried out on populations of strains ACN277, ACN293, and ADP1, and products were visualized after separation on a 1% agarose gel (d). Each PCR reaction contained 10 μ l of a template DNA stock (3 ng/ μ l) that had been diluted by the fold indicated above each lane. The two end lanes contained size markers (phage lambda DNA digested by *Hind*III).

junction identified in ACN277. This same junction sequence had been observed initially in five independently isolated amplification mutants. The difference in minimal template DNA concentration that resulted in a visible PCR signal enabled an estimation of the relative abundance of the junction in different populations. The specific tandem duplication in ACN277 should exist in all cells of this selected benzoate-grown population. Since the DNA concentration needed to be one million-fold more concentrated to yield similar results in the other strains, it was inferred that this junction was present in approximately 1 in 10^6 cells in populations of both ADP1 and ACN293 that had never been exposed to selective growth conditions. In repetitions with several independently grown cultures of ADP1 and ACN293, similar results were obtained (Table 3.1). Similar results were also obtained when the same method was used with different primer sets targeting three other recurring junctions. The abundance of each junction in unselected populations of the wild type and parent strain relative to that in the comparable mutant was 10^{-6} to 10^{-8} (Table 3.1). PCR products were cloned and sequenced to confirm that each primer pair exactly amplified the expected junction.

Engineered junction serves as a negative control and validates the PCR protocol

Additional PCR studies targeted a junction that had previously been engineered by fusing a region of DNA from the ACN277 junction (“sequence X” in fig. 3.7) upstream of *catA*. This “artificial” engineered junction, whose sequence is shown in fig. 3.6, can induce gene amplification in the parent strain by providing a substrate for homologous recombination (as depicted for the “transforming fragment” in fig. 3.2). The junction was used in the transformation assay (fig. 3.2) to generate a Ben⁺ amplified derivative of ACN293, designated ACN383 (Reams and Neidle, 2003). Although this “sequence X-*catA*” fused DNA (designated

Table 3.1 PCR analysis of the spontaneous formation of specific duplication junctions

Junction^a	Identity at duplication site (nt)^b	Amplicon size (bp)^c	Number of isolates^d	PCR product size (bp)	Relative abundance in unselected populations of <i>recA</i>⁺ strains^e	Relative abundance in unselected populations of <i>recA</i>⁻ strains^e
7774 (ACN394)	0	25,380	6	441	10 ⁻⁶ to 10 ⁻⁷	10 ⁻⁶ to 10 ⁻⁷
3760 (ACN277)	5	28,239	5	408	10 ⁻⁶ to 10 ⁻⁷	10 ⁻⁶ to 10 ⁻⁷
4620 (ACN399)	0	52,907	3	98	10 ⁻⁶ to 10 ⁻⁸	10 ⁻⁶ to 10 ⁻⁸
6127 (ACN400)	1	268,571	2	355	10 ⁻⁶ to 10 ⁻⁷	10 ⁻⁶ to 10 ⁻⁷
7820 [†] (ACN383)	6	22,177	0	135	Not Detected	Not Detected

^aThe identification number is the position in GenBank entry [AF009224](#) corresponding to the upstream endpoint of the amplicon in the *ben*-gene region. These numbers are sufficient to identify the specific duplication whose sequence is shown in fig. 3.6. Parentheses indicate representative strains with the corresponding junction.

^bNumber of identical bases in aligned sequences at the precise site of recombination (see fig. 3.6).

^cDeduced from genomic DNA sequence taking into account that the parent strain (ACN293) carries a 446 bp deletion in *catM*.

^dNumber of independently isolated spontaneous mutants with the junction.

^eThe amount of a specific duplication junction in different strains was assessed relative to the amount in an amplification mutant that contains the junction by comparing the minimal concentration of DNA required to observe a PCR product (see fig. 7 and text for further discussion). DNA was isolated from each strain to be tested using cultures never exposed to selective growth conditions. Two *recA*⁺ strains were used, the wild type (ADP1) and the parent strain (ACN293). The *recA*-disrupted derivatives of these strains (designated ADP197 and ACN487, respectively) were also tested. These four strains were all independently cultured several times and DNA samples were assayed separately. The number of independent cultures of each *recA*⁺ and each *recA*⁻ strain analyzed was five for junctions 7774 and 6127, two for junctions 4620 and 7820, and four for junction 3760.

[†]Junction 7820 was artificially constructed and used to engineer strain ACN383 (Reams and Neidle, 2003) (see fig. 3.6).

junction 7820 in Table 3.1) has the ability to confer a Ben⁺ phenotype, it was not among the 72 junctions isolated from spontaneous mutants.

To investigate whether this artificial junction could be detected in unselected populations, an appropriate primer set was designed. The primer set amplified the expected fragment from ACN383, proving that the PCR primers were functional (data not shown). However, the same primers failed to produce any visible PCR product with multiple independent populations of ADP1 or ACN293 even with the most concentrated samples of template DNA. Thus, this set of experiments demonstrated that not all possible primer sets yield a PCR product with this protocol. The inability to detect the artificial junction supports the conclusion that not all recombination events capable of conferring growth on benzoate are equally likely to occur spontaneously.

Isolation of amplification mutants depends on RecA, but tandem duplication does not

To study the RecA-dependence of forming and amplifying these junctions, we used a previously constructed *recA*-disrupted derivative of ACN293, designated ACN487 (Reams and Neidle, 2003). In the initial study, the transformation assay (fig. 3.2) was shown to depend on a functional *recA* allele (Reams and Neidle, 2003). In our current studies, we tested whether spontaneous amplification mutants could be obtained from the *recA* parent strain, ACN487, by direct selection on benzoate medium. Despite repeated attempts and prolonged incubation periods, no Ben⁺ derivatives of ACN487 were isolated. This result suggests that the generation of multiple copies of the *cat*-gene region depends on RecA.

Whether tandem duplication *per se* occurred in the absence of RecA was tested with PCR methods targeting four of the six recurring junctions. Two different *recA* mutants, ADP197 and

ACN487, were made using allelic replacement to disrupt the *recA* gene of the wild type (ADP1) and the parent strain (ACN293), respectively. When these *recA* mutants were tested for the ability to create duplications, two strategies were used to minimize the possibility that the culture inoculum already carried the targeted duplications. First, at least five copies of ADP197 and ACN487 were independently constructed. The generation and testing of multiple isolates made it unlikely that all copies would contain a specific duplication at the onset of the experiment, and it was nearly impossible that all four specific duplications would be present in five isolates of each mutant. The second strategy was to inoculate each culture with the fewest number of cells possible. A single *recA* colony was suspended in minimal medium and then diluted. A series of 10-fold dilutions were used for inoculation, and chromosomal DNA was isolated only from the culture that grew from the most dilute inoculum. Thus, it was highly unlikely that the resultant population arose from cells containing all four of the targeted junctions.

DNA samples were separately tested for the presence of all four junctions using the appropriate primer pairs and the PCR method described above. As shown in Table 3.1, the junctions tested were identified in all of the *recA* mutant strains even though none of these cultures had been grown under selective conditions. Furthermore, these recombination junctions were equally abundant in strains with and without RecA. Consistent with the inference that these site-specific duplications arise by illegitimate recombination, the data demonstrated that duplication formation is RecA-independent.

Discussion

Acinetobacter system for studying gene amplification

Gene amplification has significant implications for processes such as cancer, antibiotic resistance, bacterial virulence, bioremediation, and evolution (Hastings *et al.*, 2000; Hendrickson *et al.*, 2002; Reams and Neidle, 2003; Romero and Palacios, 1997; Stark and Wahl, 1984).

However, despite its importance, gene amplification is difficult to investigate experimentally. We used a system that exploits the natural transformability of strain ADP1 to characterize the amplicon size and copy number of 104 independently isolated mutants. In 72 of these mutants, precise amplicon endpoints were determined to analyze the sequences involved in duplication formation. To date, this study constitutes the largest scale analysis of spontaneous duplication sites.

This *Acinetobacter* system yields amplification mutants nearly exclusively. The high yield of amplification mutants stems from a selective procedure that requires increased expression of two discrete but clustered transcriptional units (*catA* and the *catBCIJFD* operon). Although this selection method exerts some limitations on the permissible upstream endpoints of the amplicons, there are no known constraints on the location of the downstream endpoints. This wide range of potential duplication sites, in combination with the transformation assay (fig. 3.2), provides a powerful method for selecting an assortment of amplification mutants. Although 58 distinct duplication sequences were determined (fig. 3.6), some of the independently isolated strains have the identical duplication junctions. Thus, recombination appears to occur at non-random locations.

Site specificity in duplication formation

Six different junction sequences were present in at least two independently isolated mutants (figs. 3.5 and 3.6). PCR methods used to study four of these recurring junctions revealed that they arise in non-selected populations (fig. 3.7, Table 3.1). Sequencing confirmed that the recombination events in the unselected populations were identical to those in the selected mutants. However, the site specificity did not rely on any obvious DNA features. There were no detectable sequence motifs or significant repeat elements (direct or inverted). Furthermore, in most of the recurring and the unique junctions, the aligned sequences that recombined had no identity at the precise recombination site (marked by vertical lines in fig. 3.6). This absence of sequence identity is characteristic of short-homology-independent illegitimate recombination (Shimizu *et al.*, 1997). In the most frequently isolated junction, found in six mutants, there was no identity at the recombination site and in an alignment of the 50 nucleotides surrounding this site there were only 8 identical nucleotides. The lack of sequence homology was notable, since site-specific recombination usually relies on common sequences in the target regions that recombine (Nash, 1996).

RecA-independent duplication formation

Consistent with the absence of sequence homology and the inferred involvement of illegitimate recombination, four of these junctions were shown to arise in *recA* mutants at the same frequency as in the presence of a functional *recA* gene (Table 3.1). In previous studies, duplications resulting from exchanges between regions lacking extensive direct-order repeats showed only weak dependence on RecA and occurred at lower frequency than recombination between regions containing large direct repeats (Franklin, 1967; Inselburg, 1967). Moreover, the

formation of deletions resulting from exchanges between short direct-order repeats of about 10 bp was reduced very little by *recA* mutation (Galitski and Roth, 1995; Inselburg, 1967).

In *E. coli*, an extensive study of duplication junctions addressed the amplification of a *lac* region on the F' episome (Whoriskey *et al.*, 1987). Sequence analysis focused on a subset of the recombination events in which one endpoint was within an approximately 400 bp region of the *lacI* gene. Interestingly, for the 30 characterized independent mutants, there were only nine different duplication junction sequences. Four of the nine junctions were isolated multiple times. Seven of the junctions, including the four recurring junctions, were characterized by short direct repeats (fewer than 15 bases) that sometimes involved broken regions of sequence identity. The two junctions for which no sequence homology was evident were each found in one isolate. Thus, these *E. coli* duplications appear to form at preferred sites, although the role of short homology and the sequence requirements for forming the duplications have not yet been determined.

A different experimental system was used to study the amplification of a genetic region of bacteriophage T4. Mutants with two to six tandem head-to-tail copies of the gene 17-18 region were derived from a phage with an altered gene 17 (Wu and Black, 1995; Wu *et al.*, 1995; Wu and Black, 1987; Wu *et al.*, 1991). Gene amplification causes hyper-production of the gene 17 protein product required for DNA packaging of the phage in *E. coli*. Three independent mutants had the identical junction site caused by recombination between two copies of a five bp sequence. The regions that undergo recombination share a 24 bp stretch of partial homology containing 4 mismatches (Wu *et al.*, 1991). It was striking that recombination in all three mutants occurred within a single 5 bp region, since the 24 bp area contains one 6 bp region of identity and another with a different 5 bp stretch of identity. Furthermore, plaque hybridization

showed that 190/191 independently isolated amplification mutants had recombined within the identical 5 base pair sequence. No other recombination events within this 24 bp region were detected. These duplications in the T4 mutants were shown to form by a sequence specific process rather than by general recombination (Wu and Black, 1995; Wu *et al.*, 1995; Wu *et al.*, 1991).

Mechanisms of duplication formation

The most striking feature of the *Acinetobacter* duplication sites is the complete absence of sequence identity at the recombination site in more than half the mutants and very little identity (1 to 2 bp) at the duplication site in most of the remaining mutants. Furthermore, most of the regions surrounding the duplication site lack short sequence repeats and regions of partial homology. The short sequence repeats in *E. coli* may be involved in illegitimate recombination mediated by end-joining reactions or strand slippage during replication (Roth *et al.*, 1996; Whoriskey *et al.*, 1987). In contrast, one role of the 24 bp region of partial homology in the T4 phage system appears to involve sequence-specific interactions with the small subunit of a two-subunit terminase used to cut and package phage DNA (Wu *et al.*, 1995). According to a proposed model, the terminase subunit helps to align the distal regions with partial sequence identity, and this alignment/synapsis facilitates illegitimate recombination catalyzed by other enzymes.

In the *Acinetobacter* mutants, the relatively large size of the amplicons and the absence of short direct repeats near most of the duplication sites argues against the participation of strand slippage during replication. Here, illegitimate recombination may involve the ligation of broken DNA ends (end joining) (Ehrlich *et al.*, 1993), but unlike many other examples, there appears to

be no involvement of short regions of homology. Topoisomerases cleave and religate DNA, and they occasionally catalyze illegitimate recombination (Ikeda *et al.*, 1982). However, topoisomerases usually religate the same strand of DNA. DNA gyrase can cleave DNA at specific sites and sometimes mediates illegitimate recombination events, possibly including the tandem amplification of a chromosomal region of *Yersinia enterocolitica* (Seoane *et al.*, 2003). Although no previously identified gyrase-cleavage sites were identified at or near the duplication junctions in our studies, target sequences can be difficult to identify and sequence specificity of the *Acinetobacter* enzyme is unknown.

Secondary or tertiary DNA structure, including nucleoid arrangements, could facilitate site-specific DNA end-joining reactions catalyzed by DNA gyrase or other enzymes. Interestingly, a large portion of the mutants (45/104) amplified a nearly 300 kbp chromosomal segment whose downstream end was within the *pca* genes. Portions of the supraoperonic clusters encompassing the *pca* and *cat* genes share significant sequence homology and similar functions in aromatic compound degradation (Harwood and Parales, 1996; Young *et al.*, 2001). For example, *catIJF* and *pcaIJF*, which encode isozymes, are 98% identical in nucleotide sequence. Although we failed to identify any relationship between the relative positions of strong sequence homology and the duplication sites, the high frequency of recombination between these regions suggests some type of physical interaction. These interactions could be promoted by the structural organization of DNA within the cell. Sequence homology, which anchors RecA-dependent strand transfer in this *Acinetobacter* strain has the ability to facilitate illegitimate recombination elsewhere on the same DNA molecule (de Vries and Wackernagel, 2002). However, in our studies, disruption of *recA* did not alter the frequency with which duplication occurred at a site in the *pca* region (junction position 6127, Table 3.1).

Not all of the junctions in this study need to have been generated by the same mechanism. Certainly, the two junctions with IS/236 (fig. 3.6) are likely to have formed by homologous recombination between one IS/236 element in the *ben* region and another in the downstream region. Presumably, insertion of this element into the upstream and downstream sites occurred prior to duplication formation. Furthermore, naturally occurring direct repeats in the ADP1 genome most likely cause additional duplications. However, the absence of such sequences in the regions that confer the selected phenotype may have skewed the types of mutants that were isolated in this study.

Two step amplification process

The Ben⁺ amplification mutants were not obtained from a parent strain lacking *recA*, even though the appropriate duplications form in this strain (Table 3.1). These results implicate a two-step amplification mechanism in which the first step (duplication via illegitimate recombination) is followed by RecA-mediated amplification of a region with copies of identical sequences (Romero and Palacios, 1997; Roth *et al.*, 1996). Presumably, the initial duplication creates a substrate for homologous recombination (fig. 3.1), which allows further amplification over the course of many cell generations by repeated cycles of sister strand exchange during replication.

Since colony formation appears to depend on the RecA-mediated amplification of a duplicated region, it would be of interest to know how many cells in the starting population carry an appropriate duplication. Previous bacterial studies have shown that tandem duplications of most chromosomal loci occur frequently (Roth *et al.*, 1996). In our studies, it is difficult to estimate the number of different spontaneously occurring duplications that encompass the *catA*

gene and *catBCIJFD* operon while excluding the promoter mutation of *benA*. Furthermore, whether all such duplications occur at the same frequency is unclear. Several individual site-specific duplications that meet these criteria each appear to be present in an unselected population of 10^6 to 10^8 cells (Table 3.1). Thus, there should be a sufficient number of cells carrying duplications in the initial population to account for the observed frequency of approximately 10^{-8} amplification mutants after 3 weeks of incubation. During the incubation period, cells with multiple copies of the *cat* genes should have a selective growth advantage. As the copy number of the *cat*-region amplicon increases, the cells will grow faster. A process of natural selection should continue until a population of amplification mutants is visible as a colony on the benzoate medium plate.

It has previously been noted that most cells in a population of amplification mutants grown in liquid benzoate medium attain an optimal copy number characteristic of the specific amplicon (Reams and Neidle, 2003). Natural selection appears to enrich for cells whose copy number optimizes the growth rate for the selective condition. The amplification mutants in this study confirm that the optimal amplicon copy number tends to be higher when the amplicon size is small (fig. 3.4) and that mutants grow more quickly when copy number is higher (unpublished data). However, there was no correlation between amplicon size and the initial time it took for a colony to form during selection on a benzoate-medium plate (unpublished data). Such apparent variability in the time needed for colony formation may reflect that the RecA-dependent amplification of a duplicated region within any given cell can occur in a stochastic fashion at different times during the incubation period.

Evolutionary implications

The significance of gene duplication for the evolution of individual genes has long been accepted (Ohno, 1970). A related topic is how the co-amplification of functionally related genes affects the evolution of genome organization. In this *Acinetobacter* study, a selective advantage was only conferred to a mutant if two distinct but functionally related transcriptional units were co-amplified. Nevertheless, such mutants were readily isolated. This evidence that the spontaneous co-amplification of multiple transcriptional units can confer selective advantage has important evolutionary implications for the clustering of functionally related genes.

Genome analysis has been drawing recent attention to gene organization at a higher level than operons. Yet, the tendency for functionally related bacterial genes and operons to be organized in supraoperonic clusters was first noted more than thirty years ago (Rosenberg and Hegeman, 1969; Wheelis and Stanier, 1970). New terms such as genomic island and uber-operon, which describe variations on the clustering theme (Hacker and Kaper, 2000; Lathe *et al.*, 2000), emphasize that genes are not dispersed randomly during evolution but are maintained within a genome in a functional context. This organization cannot be justified by transcriptional requirements alone, and various theories have been proposed to explain the phenomenon. Horizontal gene transfer may play a role (Lawrence, 2003).

An additional possibility is that gene amplification contributes significantly to the advantage of having functionally related genes in close proximity. Such a genetic arrangement provides the opportunity for a complex metabolic function to be regulated by gene dosage (Roth *et al.*, 1996) (as in this study). Furthermore, the function can evolve after amplification in ways not possible if the associated genes were unlinked. Amplification allows mutations to accrue within the genetic region while at least one copy of each of the genes maintains functionality.

De-amplification can subsequently reduce or eliminate homologous segments, allowing the most advantageous genotypes to be selected (Hendrickson *et al.*, 2002). The theory applied to the duplication, divergence and evolution of an individual gene applies to a group of functionally related genes only if they are co-amplified. These *Acinetobacter* studies provide dramatic experimental evidence that co-amplification does generate mutants whose selection depends on a clustered gene organization. This evidence supports the theory that gene amplification is a driving force in the chromosomal clustering of functionally related genes.

Experimental Procedures

Bacterial strains, growth conditions, and the selection of Ben⁺ mutants

Acinetobacter strains were derived from the wild-type ADP1 (also designated BD413) (Juni and Janik, 1969). Several mutants were previously described (ADP197, ACN277, ACN282-ACN284, ACN293, ACN383, and ACN487) (Gregg-Jolly and Ornston, 1994; Reams and Neidle, 2003). Recombinant plasmids were maintained in *E. coli* DH5 α or Top10 (Invitrogen) (Grant *et al.*, 1990). Except where noted otherwise, bacteria were cultured at 37°C in Luria-Bertani medium or minimal medium with 10 mM succinate or 2 mM benzoate as the carbon source (Sambrook *et al.*, 1989; Shanley *et al.*, 1986). When needed, antibiotics were added at the following final concentrations: kanamycin, 40 μ g/ml; streptomycin, 13 μ g/ml; spectinomycin, 13 μ g/ml.

To isolate spontaneous Ben⁺ mutants, single succinate-grown colonies of strain ACN293 were taken from separate plates and each was used to inoculate separate batches of liquid succinate medium. These cultures were individually plated on minimal medium with benzoate as the sole carbon source, as described previously (Reams and Neidle, 2003). Approximately 5

$\times 10^9$ ACN293 cells were spread per plate. Mutant colonies arose during incubation at room temperature (approximately 22 °C) and were chosen for further analysis. No more than one Ben⁺ mutant was selected from each parent culture. The phenotypes corresponding to the terms Ben⁺ and Ben⁻ were defined previously to clarify the designations used for cells that grow very slowly on benzoate (Reams and Neidle, 2003).

Assessment of amplicon size and copy number by pulsed-field gel electrophoresis and hybridization methods

A previously devised method was used to determine the size of amplicons with a type of pulsed-field gel electrophoresis, transverse alternating field electrophoresis (TAFE, Geneline II, Beckman) (Reams and Neidle, 2003). The entire genome of each mutant was analyzed. Briefly, this method exploits the relative positions of six wild-type *NotI* restriction sites. In the mutants, amplification either generates a new *NotI* fragment of the exact size as the amplicon or it increases the size of one fragment (designated E in fig. 3.4) by an amount equal to the amplicon size multiplied by the number of copies present (Reams and Neidle, 2003). In all cases, the estimation of amplicon size by this approach was consistent with that deduced by comparing the DNA sequence at the boundaries of the amplicon to the genome sequence of the wild-type strain (http://www.genoscope.cns.fr/externe/English/Projets/Projet_DY/DY.html). Southern hybridization methods were carried out as previously reported with nonradioactive probes (Gralton *et al.*, 1997). To determine amplicon copy number, the hybridization strengths of single-copy fragments were compared to those of amplified regions. Hybridization strength was assessed by digital analysis of signals on exposed X-ray films (Reams and Neidle, 2003).

Generation of plasmid libraries and transformation assay to detect plasmid-borne mutant DNA that confers a Ben⁺ phenotype to ACN293

To make plasmid libraries, chromosomal DNA from each mutant was purified, digested with either *EcoRI* or *XbaI*, and ligated to the cloning vector pZero-2.2 (Invitrogen) as previously described (Reams and Neidle, 2003). This cloning vector (a ColE1-type replicon) is designed for the selection of kanamycin-resistant *E. coli* host cells that all carry recombinant plasmids. To identify DNA fragments from the mutants responsible for growth on benzoate in the absence of the normal *Acinetobacter* transcriptional regulators, a previously described transformation assay was used with ACN293 as the recipient strain (Reams and Neidle, 2003). In an initial screen, individual *E. coli* cells that carried recombinant plasmids were patched to a benzoate medium plate on which the recipient strain had already been spread. Growth in the vicinity of a patched colony indicates that *Acinetobacter* DNA released from the *E. coli* donor colony can transform ACN293 to a Ben⁺ phenotype (Reams and Neidle, 2003). Since the cloning vector used is not stably maintained in ADP1, growth should result from incorporation of the mutant DNA into the recipient by homologous recombination rather than from true complementation. To confirm positive results in the initial screen (fig. 3.2), the corresponding plasmid was purified and linearized by digestion with an appropriate restriction enzyme and used as donor DNA in additional transformation tests (Reams and Neidle, 2003).

Sequence analysis of duplication sites

Recombinant plasmids with DNA capable of transforming ACN293 to a Ben⁺ phenotype were analyzed by restriction digest patterns to devise a DNA sequencing strategy. Plasmid sequencing was carried out at a University of Georgia Integrated Biotechnology Laboratories with universal

pUC/M13 sequencing primers (Promega). Additional sequencing reactions were performed with specifically designed primers purchased from Genosys Biotechnologies. In all cases the sequence of the duplicated DNA exactly matched a region of the wild-type genome.

PCR assays to detect duplication junctions

To study the relative abundance of specific duplication junctions in different *Acinetobacter* populations, chromosomal DNA was isolated by a phenol-extraction based method (Sambrook *et al.*, 1989). DNA was extracted from cultures grown in minimal medium with succinate (for unselected populations) or benzoate (for selected mutants). To test different populations of the same strain, independent colonies were used to inoculate multiple cultures that were each assayed by the PCR method. For repetitions with the same strain, colonies were taken from separate plates that had been streaked on different days.

Prior to use as template in PCR reactions, a sample of the chromosomal DNA was digested with *EcoRI*. After digestion for 1 hr, the restriction enzyme was heat-inactivated at 68°C for 15 min. DNA concentrations were determined using a fluorometer (TD-360, Turner Designs), and each sample of *EcoRI*-digested chromosomal DNA was standardized to 3 ng/μl. This DNA was subjected to a series of 10-fold dilutions. Multiple individual PCR reactions, each in a 50 μl total volume, were performed with DNA in amounts ranging from 30 ng to 3 x 10⁻⁹ ng. In addition to template DNA, each PCR reaction contained the following reagents: 0.5 μM of each primer, 100 μM of each dNTP, 1X Faststart Taq buffer, and 1 unit Faststart Taq polymerase (Roche). The PCR reactions, performed in a thermocycler (MJ Research MiniCycler with a hot bonnet), were initiated by heating to 95°C for 4 min. The remainder of the 40-cycle

reaction was conducted at 94°C for 45 seconds (denaturation), 55°C for 45 sec (annealing), 72°C for 30 sec (elongation) followed by a 7 min incubation at 72°C to complete the synthesis.

The following primers were used in the PCR to detect 5 different duplication sites (see Table 1 for the junction designations): 3760up103 5' GATCCGCGGCACATGTCGT 3' and 3760down305 5' CATGGTTGCGCGGTAAATAT 3' (junction 3760), 7774up362 5' TATTTTCGATAGTCGTTTCTA 3' and 7774down79 5' GAGCAAGCGAGTGGAGGAG 3' (junction 7774), 4620up 5' CGGGGTAACGCCATCTTCAAATTG 3' and 4620down 5' GGAGGGATCAAACCTTGGTACTGG 3' (junction 4620), 6127up102 5' TTAGTTCCTCCCGTTGCAA 3' and 6127down253 5' AGGGACTGGCGGCTTCACA 3' (junction 6127), 3760up2 5' AGTATTGATGCTGTGCCA 3' and newseq 5' CTCAGGCCAATGATAAAC 3' (junction 3760), 3760up2 5' AGTATTGATGCTGTGCCA 3' and catAup 5' GCAGATATCGCTAAATACCTTC 3' (junction 7820). The PCR products were separated on a 1% agarose gel using standard electrophoresis procedures. The presence or absence of a PCR product was determined by eye after visualization by ethidium bromide staining. PCR products cloned, with the pCR2.1-topo vector (Invitrogen), and sequenced at a University of Georgia facility (the Integrated Biotechnology Laboratories).

Disruption of the chromosomal copy of recA in the wild-type and parent strains

Although the two *recA*-disrupted strains (ACN487 and ADP197) had previously been made (Gregg-Jolly and Ornston, 1994; Reams and Neidle, 2003), several independent versions of these strains were reconstructed. As described above, new strain versions were generated because of duplications that might be present in a given mutant. Previously described allelic replacement techniques (Gralton *et al.*, 1997) were used in which independent cultures of the parent strain

(ACN293) or the wild-type strain (ADP1) were used as the recipients in transformations to generate ACN487 or ADP197, respectively. A linearized version of a plasmid (pZR106) containing a Tn5-disrupted allele of *recA* was used to transform the recipient (Gregg-Jolly and Ornston, 1994). Selection of the appropriate drug marker was used to isolate the corresponding strain in which this allele had replaced the wild-type *recA*. Correct genotypes were confirmed by Southern hybridization methods (Gralton *et al.*, 1997).

Acknowledgements

We thank Brian Higgins, Dr. Anna Karls, and Dr. John Roth for helpful discussions. This study was supported by National Science Foundation grant MCB 0212604 to E.L.N.

CHAPTER 4

GENE AMPLIFICATION IN *ACINETOBACTER*:

EVALUATION OF ADAPTIVE MUTABILITY¹

¹Reams, A.B. and E.L. Neidle. 2003. To be submitted to *PNAS*.

Summary

In different experimental model systems, amplification mutant colonies arise on growth-limiting selective medium after a significant delay. A new theory termed “adaptive amplification” states that this delay reflects an amplification mechanism that is inducible by stress caused by the selective environment. It has been suggested that this stress most likely induces the formation of a tandem duplication that can rapidly become further amplified. Controversy surrounds the “adaptive” ability of the environment to cause genetic variation during evolution. In this report, we used a selection for *Acinetobacter* gene amplification mutants to test adaptive amplification. Here we show that duplications arise spontaneously and prior to selection. Furthermore, generation from two copies to many copies may take several weeks during the original selection. Based on these results, the delay in colony formation is explained by a natural selection process whereby higher amplification from preexisting duplications occurs slowly via a series of homologous recombination events. This implies that mutant colony formation occurs by a selective process and that duplication formation is not adaptive. However, gene amplification was shown to be inducible by DNA-damaging agents. These agents may affect a pathway responsible for amplifying preexisting duplications. Finally, the natural transformability of *Acinetobacter* was shown not to be required for gene amplification.

Introduction

The role of the environment in inducing gene amplification has implications for significant biological processes such as cancer and for understanding evolution. A theory termed “adaptive amplification” states that this type of genetic change is stimulated in response to stressful and growth-limiting environmental conditions (Hastings *et al.*, 2000). Currently, there is

considerable debate concerning the ability of the environment to induce mutations, to direct mutations to advantageous sites, and to alter the rate of evolution. To address the topic of stress-induced mutation experimentally requires methods for determining the occurrence of the underlying genetic change(s) relative to the onset of the non-lethal stressful condition. Such measurements are indirect, and experimental results have been subject to alternative interpretations. In this report, we evaluate whether the spontaneous amplification of a chromosomal region involved in benzoate degradation by the bacterium *Acinetobacter* sp. strain ADP1 occurs adaptively. Gene amplification, observed in hosts ranging from viruses to humans, can result in cancer, increased virulence and antibiotic resistance. These consequences underscore the importance of understanding factors that induce gene amplification.

Most studies of adaptive processes focus on point mutation, often using methods that assess the reversion of a *lac* allele frameshift mutation in bacteria. However, it has been argued that gene amplification in the *lac* frameshift system can underlie the appearance of adaptive point mutation without normal mutation rates being induced directly by starvation (Andersson *et al.*, 1998). This argument was countered by another study which led to the conclusion that gene amplification constitutes an adaptive mechanism distinct from those involved in point mutation (Hastings *et al.*, 2000). Furthermore, these authors propose that the formation of the original tandem duplication via non-homologous recombination is most likely the stress-induced step, and once a tandem duplication is generated subsequent higher amplification occurs rapidly. Our studies do not address the disputed role of amplification in *lac* reversion. Instead, our recently developed method for analyzing spontaneous chromosomal alteration in a soil bacterium provides an ideal system for testing the adaptive nature of gene amplification.

This *Acinetobacter* system uses a parent strain lacking two transcriptional activators that normally regulate genes for benzoate catabolism in the wild-type strain ADP1. In amplification mutants, multiple copies of a region encompassing two chromosomally-clustered weakly expressed transcriptional units compensate for the loss of transcriptional activation. After inoculation with the parent strain, selection for growth on benzoate as the sole carbon source yields amplification mutants nearly exclusively. For example, 104 of 105 mutants recently isolated by this method each had an amplified chromosomal segment (amplicon) ranging in size from approximately 12 to 290 kbp and in copy number from 3 to 30. Furthermore, the precise boundaries of each amplicon can be determined using a transformation assay based on the natural competency for DNA uptake by these *Acinetobacter* strains.

Previous studies indicate that spontaneous chromosomal amplification in ADP1 is a two-step process involving first a gene duplication event that usually occurs via RecA-independent illegitimate recombination. Most of the duplication sites lack evident sequence identity (including short regions of homology) between the chromosomal regions that recombine. Nevertheless, at least several duplication events are detectable in independent unselected populations of the wild type and parent strains as well as in *recA*-disrupted derivatives of both strains. While these duplications were evident in populations of the *recA*-disrupted parent strain, they did not yield colonies (designated Ben⁺) when selected on medium with benzoate as the carbon source. Thus, the second step of the amplification process appears to depend on homologous recombination between the duplicated copies of the amplicon to increase the copy number above two.

The key question addressed in the current investigation is whether either step in the *Acinetobacter* amplification process occurs in response to the growth-limiting selection

conditions. To do this, we tested whether each of these two steps of amplification were adaptive by determining whether they occurred before or after selection. Another goal was to test whether gene amplification was inducible by DNA-damaging treatments and agents, which are known to induce stress responses. Based on our results, a schematic model for the pathway of gene amplification is proposed.

Results and Discussion

Amplified Ben⁺ colonies originally arise after a prolonged delay

Using a previously described method, spontaneous amplification mutant colonies that grow on benzoate (Ben⁺) were selected from a parent strain, ACN293. The parent strain fails to grow on benzoate because it lacks the transcriptional activators normally used to express high levels of the *cat* genes (Reams and Neidle, 2003). This selection is nearly exclusive for mutants carrying gene amplification and mutant colonies arose after a significant delay. Approximately 5×10^9 ACN293 cells were spread per plate with benzoate as the sole carbon source, and after prolonged incubation the number of Ben⁺ colonies increased (fig. 3.3). For example, at the end of nine days of incubation approximately one colony on average was present on a given plate. At the end of the fourth week, there were approximately 100 colonies per plate. This result corresponds to colony formation at a frequency of approximately 10^{-8} after 3 weeks of incubation. Amplified mutants formed visible colonies as early as 3 days and additional colonies continued to arise over time until the plates eventually dried out after about 50 days. The delay in colony formation was not a direct effect of the slow growth of the amplified mutants since when cells were directly transferred from their original amplification colony to fresh selective plates they formed large single colonies similar to the growth rate of the wild type, within 1-2 days. On the contrary, this

delay in colony formation has been reported previously in high copy amplification mutants and is the basis behind the theory of adaptive amplification (Hastings *et al.*, 2000). According to this theory, this delay reflects amplification having a mechanism inducible by the stress of the growth-limiting selective medium.

One of the advantages of the *Acinetobacter* system over many other amplification systems to test the theory of adaptive amplification is that it is a relatively exclusive selection for mutants carrying gene amplification. In analysis of 105 spontaneous Ben⁺ mutants, all but one of these mutants carried tandem head-to-tail repeats of a chromosomal segment. Only a single Ben⁺ colony was found which did not carry amplification (Reams and Neidle, chapter 3 of this dissertation). The single Ben⁺ mutant with no evident chromosomal rearrangement arose after two days of incubation, whereas all amplification mutants initially required at least three days to form a Ben⁺ colony. No other Ben⁺ colonies arose within just two days incubation. Although the rate at which such a non-amplification Ben⁺ mutant arises is unknown, it appears to be extremely rare (<1%) in this system as it has been detected in only one case during our studies.

Rate of visible colony formation does not necessarily reflect amplicon length

Amplification mutant colonies arise gradually over a wide span of incubation time, ranging from as little as 3 up to 50 days (fig. 3.3). One possibility to explain the reason amplified isolates show such a wide variety in the duration before they form visible colonies could be that growth rate is reflective of the amplicon size. Consistently, in *E. coli lac* gene amplification mutants, Hastings and co-workers observed the variation in colony emergence can be explained by a direct correlation between amplicon length with the number of days required for colony formation (Hastings *et al.*, 2000). The greater the amplicon size, the more days are required for

visible colony formation. It was suggested this may reflect the burden of replicating the extra DNA or from increased instability of the longer arrays.

In the *Acinetobacter* amplification mutants, there was not a consistent correlation between amplicon size and the number of days required for colony formation (fig. 4.1). On one hand, we found that a few colonies with smaller size amplicons (<100 kb) arose prior to 11 days while those with larger size amplicons (>100 kb) require a minimal of 11 days incubation time (fig. 4.1). However, after this period, amplicon size was not necessarily attributed to slower colony formation. After 11 days incubation, mutants with wide varieties of amplicon size formed visible colonies after the same incubation time. For example, three independent amplification mutants arising on the 11th day had amplicon sizes of approximately 32, 146, and 273 kb, respectively. On the 25th day, two amplification isolates had amplicon sizes of 53 and 274 kb, respectively. Each of these colonies were not sister colonies derived from previously arisen colonies on the same plate since all of these mutants were the first colony to arise on the plate. Therefore, in general amplicon size alone does not sufficiently explain the delay in colony formation in the *Acinetobacter* system.

These differences in observations between the *lac* system studied by Hasting and colleagues and the one described here may merely reflect differences in the range of data analyzed. In the *E. coli* system, all 11 of the total amplification mutants analyzed arose between 2-6 days. Most of the *Acinetobacter* amplification mutants arose after this period, with only four amplification mutants arising between 2-6 days. Interestingly, one of these four amplification mutants appeared on day 6 and had an amplicon size smaller (12 kb) than the three amplification mutants arising after just 3 days (19-26 kb). In addition, all 11 amplification mutants analyzed in the *lac* system had amplicon sizes less than 50 kb (Hastings *et al.*, 2000). Amplicon sizes from

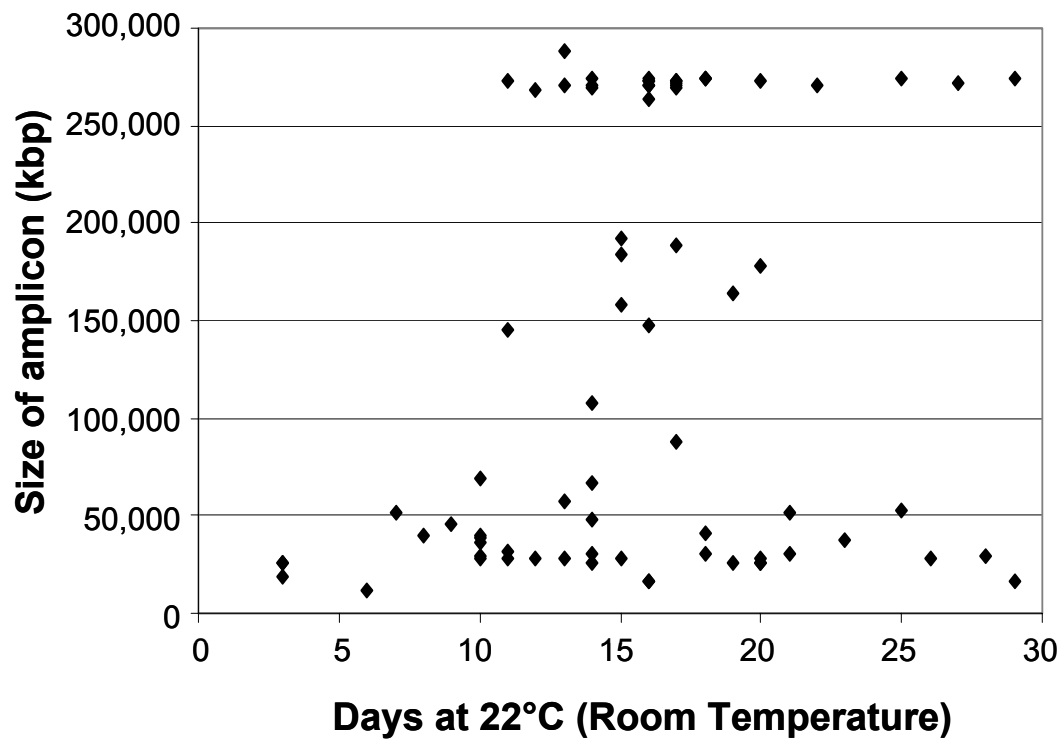


Fig. 4.1 Graph showing the correlation between amplicon size and the number of days until visible colony formation.

The X-axis represents the number days until visible colony formation on the original benzoate plate from which the amplification mutant was selected.

104 *cat* gene amplification mutants ranged from approximately 12 up to 290 kb. When amplification mutants were categorized into groups based on their amplicon size, it was shown that some amplicon size groups were more common than others (Reams and Neidle, please see chapter 3 of this dissertation). Still, when mutants were analyzed within each amplicon size group the day of colony formation varied significantly. For example, colonies with amplicon sizes ranging between approximately 12-50 kb arose anywhere between day 3 and day 29. Therefore, there does not seem to be a direct correlation between amplicon length and day of colony formation in the *Acinetobacter* system.

Testing whether the first step of amplification is adaptive

In order to test whether the initial duplication step of amplification was adaptive, we attempted to detect whether *cat* gene tandem duplications arose prior to selection. In a previous study, PCR primers were designed to amplify over four of the site-specific duplication junctions. These primers were designed in an orientation that a PCR product should only be produced if this specific duplication was present in a population. These products were further verified by examining their size on an agarose gel as well as sequencing over the cloned PCR fragment. Using this method, each of the four site-specific duplication junctions that were examined were detected in all of the tested independent populations grown under nonselective conditions. These four strains were ACN293, ADP1, and *recA*-disrupted derivatives of both these strains (Reams and Neidle, please see chapter 3 of this dissertation). This observation raises an important point: if the second step of amplification occurs rapidly under the original selective conditions and if these four duplications are present prior to selection, then amplification mutant colonies carrying these junctions should arise rather quickly during selection. However, the independently isolated

mutants carrying any of these four junctions arose on different days over a broad range of time (table 4.1). This evidence suggests that cells carrying a duplication may take several days to generate a visible colony under the original selection.

To test further whether duplications arose prior to selection, a unique junction identified from a gene amplification mutant was attempted to be detected in the original unselected culture from which it arose. The junction analyzed was isolated from an amplified Ben⁺ strain whose colony was originally visible after 16 days on selective media. The amplicon size was determined to be 16,086 bp in length based on comparing the junction's end points with the available genome sequence (www.genoscope.com). This size was consistent with PFGE and Southern hybridization analysis. Using the PCR method, primers were designed specifically for this junction and it was detected at a very low frequency ($\sim 10^{-8}$) in the chromosomal DNA of the original unselected culture from which it arose (data not shown). Sequencing of the cloned PCR product confirmed the specific junction targeted. This reaffirms that cells carrying *cat* gene duplications can arise prior to exposure to selective conditions. Furthermore, while it can not be proven that the cell(s) carrying this duplication inevitably gave rise to the amplification mutant colony characteristic of this junction, it further suggests that the process of generating higher amplification from a preexisting tandem duplication may take several days on the original selective medium (i.e. 16 days in this case).

Isolation of derivative strains with various copy numbers

In order to investigate the kinetics of the second step of amplification, where higher amplification is generated from an initial duplication, a derivative strain was isolated carrying a single tandem duplication. Using a previously described method, an amplification strain

Table 4.1 Independently isolated mutants with identical junctions arise on different days.

Junction*	Size of amplicon	Number of mutants	# days of original colony emergence
7774	25,380 bp	6	3, 3, 14, 19, 20, 20
3760	28,239 bp	5	12, 13, 14, 15, 20
4620	52,907 bp	3	25, 36, 43
6127	268,571 bp	2	12, 36
7714	270,135 bp	2	16, 33
6551	273,077 bp	2	17, 20

* Jct number refers to the GeneBank number within the *ben* region corresponding to the upstream endpoint of the amplicon. The first four junctions (7774, 3760, 4620, and 6127) were previously analyzed and detected in multiple independent populations grown under nonselective conditions. These nonselective populations were comprised of the following four strains: ACN293, ADP1, and *recA*-disrupted derivatives of both these strains (Reams and Neidle, please see chapter 3 of this dissertation).

carrying high-level tandem amplification (~20 copies) of an approximately 28 kb amplicon was grown under non-selective conditions, succinate (Reams and Neidle, 2003). Such growth conditions do not require amplification for growth and continual growth was accompanied by deamplification. After 200 generations of nonselective growth, derivative strains were isolated that carried various copy numbers of the identical amplicon. A frozen stock of each of these derivative strains was made and grown again on succinate medium to obtain sufficient cells for DNA isolation. This procedure corresponded to another 30 generations of growth on non-selective medium. To determine the amplicon copy number of each strain, plugs were prepared and analyzed by PFGE and Southern hybridization. Figure 4.2 shows the results of three of these derivative strains varying based on amplicon copy number. Each of these three strains appeared as “frozen” intermediate stages of in the process of deamplification. The majority of the cells in populations of these strains contained two, three, and eight copies of the 28 kb amplicon, respectively. Within each population, there was some variation in amplicon copy number. This is most evident in the strain with eight copies by the ladder of fragments below the intense band. Nonetheless, the majority of the cells in each population appeared to contain a fixed copy number. In general, the lower the average number of copies, the longer the population took to grow when transferred back onto selective medium. Further analysis by PFGE and Southern hybridization of these benzoate-grown derivative strains demonstrated that they had regenerated higher amplification during selective growth. In contrast, strains which reverted completely back to a haploid state were not able to grow on benzoate medium nor did they generate higher amplification.

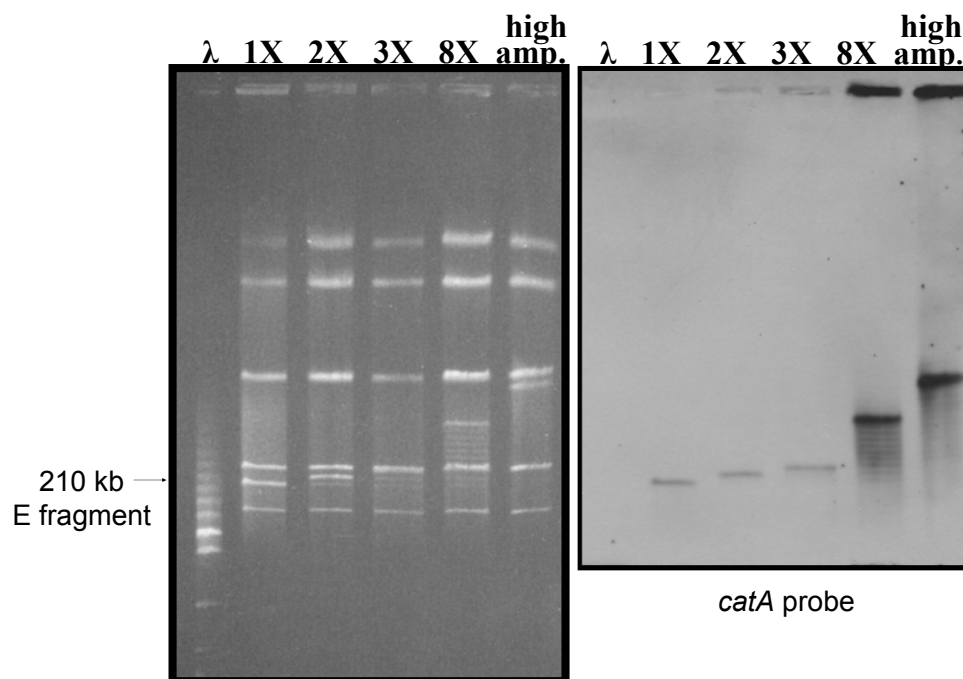


Fig. 4.2 Derivative strains with different levels of amplification.

The Pulse-Field Gel of these derivative strains is on the left while the Southern hybridization of this gel is on the right. Each derivative strain was digested with *NotI*. The *cat* genes are located on the *NotI* “E-fragment”. A *catA* probe was used for the Southern hybridization. A λ ladder in the far left lane was used as a size marker.

Test whether the second step of amplification is adaptive

The claim that amplification is adaptive is based on evidence provided by reconstruction experiments (Hastings *et al.*, 2000). The aim of these experiments is to determine whether amplification (both steps) occurred before or after exposure to selection. In these experiments, cells carrying high amplification are diluted and plated along with a high density of parent strain cells onto selective media. These conditions mimic the original selective conditions in which the amplification mutants initially arose. The high density of parent strain cells work as scavengers which compete with cells containing preexisting amplification for available carbon source. This competition is growth restricting and may be the cause for the delay in colony formation. Under these conditions, if the supplemented amplification mutants form visible colonies in less time than it took for the original colony to form, then amplification is presumed to have generated during the stress inducted by the selection and is thereby considered adaptive. Indeed, Hastings and co-workers found that amplification mutants arose in fewer days in the reconstruction experiment than the number of days before these amplification colonies were visible on the original plate on which they arose. While, these finding did not distinguish which of the two steps of amplification is adaptive, these authors proposed that the cause for the delay was the formation of the original duplication. This latter theory presupposes that once a duplication is generated (via non-homologous recombination), higher amplification can occur relatively rapidly under the original selective conditions.

To test whether amplification was adaptive in *Acinetobacter*, we also performed these reconstruction experiments with the *cat* gene amplification mutants. Furthermore, to test whether the second step of amplification was the cause for the delay in colony formation, high amplification cells were substituted with the derivative strains carrying different levels of

preexisting amplification (fig 4.2). Since these lower copy number strains generate higher amplification when grown on benzoate, the reconstruction experiments were performed with these derivative strains to investigate the length of time necessary to generate higher amplification under conditions which mimic the original selective conditions. The results from these experiments are shown in figure 4.3. Where the amplification mutant colonies initially arose after 12 days on the original selective media, most of the cells carrying preexisting amplification arise within only four days under the conditions of the reconstruction experiments. Based on these results, it was concluded that higher amplification was generated under the original selective conditions. In the reconstruction experiments with the derivative strains, the results were significantly different. The lower the initial copy number, the longer the average number of days these cells took to form colonies. For example, most cells with only two copies took at least 27 days to form a visible colony under the original selective conditions. PCR analysis on chromosomal DNA from 30 Ben⁺ colonies arisen after 20 days incubation derived from the reconstruction experiments with the two copy strains, verified that 12 of these 30 colonies contained the same duplication junction present as the strain containing the preexisting two copies. The other 18 colonies were considered Ben⁺ mutants derived from the scavenger parent strain population. PFGE and Southern hybridization analysis on these 12 strains showed a ladder of bands hybridizing to a *catA* probe indicating that all of these colonies had regenerated higher amplification from the preexisting duplication. Based on these observations, we conclude that the generation of higher amplification from a duplication is the cause for the delay in visible colony formation under the original selective conditions.

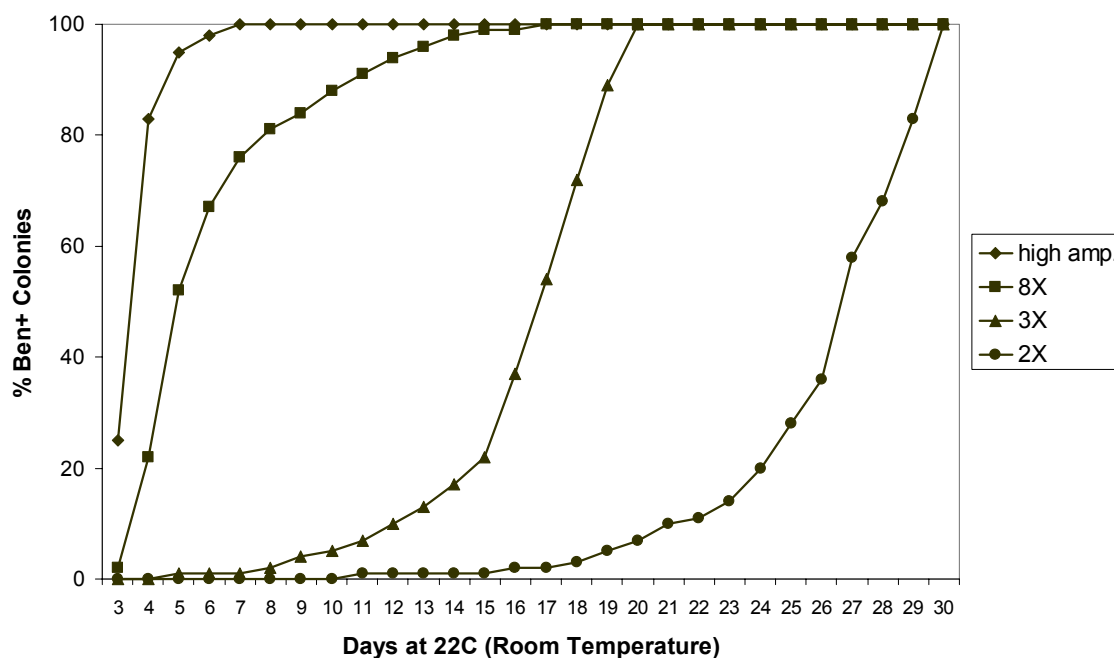


Fig. 4.3 Reconstruction experiments with strains carrying various levels of amplification. Derivative strains carrying different copy numbers of an identical amplicon were obtained from the deamplification experiments described in chapter 2. The derivative strains were diluted at 10^{-8} with the parent strain and spread plated onto minimal media plates with benzoate as the sole carbon source. The “high amp.” refers to a strain carrying approximately 15-copies of a 30 kb amplicon containing the *cat* genes. The 8X, 3X, and 2X strains carry 8, 3, and 2 copies of this amplicon, respectively.

The delay in colony formation can be explained by a natural selective process rather than an adaptive process

According to the theory of adaptive amplification, since the second step of amplification is only apparent after exposure to selection, this step could be considered to be adaptive. However, we believe this second step need not necessarily be induced by the stress of the selective media, but rather it may be explained by a natural selective process. In this selective process, a rare occurring event which increases the copy number is selected within a clonal population carrying a preexisting duplication. This rare occurring event may not be induced but rather selected based on the presence of a larger number of gene copies which contribute to increased fitness under the selective environment. The reason high amplification is not apparent prior to selection is most likely due to the instability of having high copy number of a long homologous repeats.

The delay in colony formation may be explained by a natural selection process whereby higher amplification from preexisting duplications was slowly generated through a selected series of homologous recombination events. This process can be explained by a diagram dissecting a Ben⁺ colony into a heterogeneous population varying on amplicon copy number as proposed by Hendrickson and colleagues (2002). In this model, a population of cells (approximately 5×10^9) is spread onto a selective medium plate. Several cells within the population generate a duplication of the *cat* genes during the pregrowth phase. These cells carrying a duplication grow very slowly due to competition among approximately five billion single copy cells for the available carbon source. While the cells with two copies of the *cat* genes may grow twice as fast than those with only a single copy, they still grow very poorly and do not form a visible colony. Evidence to support slow growth of cells with fewer copies on selective medium is provided by the growth phenotype of a single copy strain and a *recA* mutant

strain with two copies, which do not amplify on selective media. These strains grow very slowly and do not form large colonies even when they are spread out on a benzoate plate and don't have neighboring cells to compete for carbon source (see below for further discussion). Under the original selective conditions, at some point within the small colony of cells with *cat* gene duplications one cell may undergo a rare occurring event that increases its copy number (i.e. from two copies to three). This event does not need to be stress-induced since the increase allows these putative clones within the emerging colony to outgrow those with fewer copies. This process may occur several times within the population until those cells with higher copy numbers outgrow the remainder of the colony. The earlier the rare event increasing copy number occurs, the faster a visible colony may form. Evidence supporting this series of events is when plugs are prepared from Ben⁺ cells immediately taken from the original selective plate their colony arose, PFGE analysis reveals a ladder of fragments representative of a mixed population of cells varying on copy number. These ladders have a significant portion of the cells with lower copy number. This ladder may reflect multiple-step events that take place during the generation of higher amplification. If these amplified mutants are transferred several times to fresh selective medium, the population will predominately consist of cells with high amplification, and no cells with low copy number will be detected. This also appears to reflect the selective nature of growth on benzoate. It has been shown previously that cells acquire an optimal amplicon copy number after prolonged growth on benzoate. Based on these results, amplification in this *Acinetobacter* system has not necessarily been a response to stress; rather it can be explained by a selective process.

RecA is required for the second step of amplification, but not the first

Our next goal was to attempt to identify genes involved in the second step of amplification. In bacterial systems, it has been widely reported that RecA is involved in the gene amplification mechanism (Roth *et al.*, 1996, Romero and Palacios, 1997). Consistently, a *recA* mutant of the *Acinetobacter* parent strain does not form Ben⁺ colonies (Reams and Neidle, 2003). However, it remains unclear at which step RecA is involved. It was previously shown that the initial duplication in *Acinetobacter* amplification mutants resulted from an illegitimate recombination event (Reams and Neidle, unpublished results). Furthermore, at least some of these events were shown to be RecA-independent.

The next goal was to test whether RecA was required for the second step of amplification. To do this a *recA* mutant strain carrying two copies of the amplicon was constructed and tested for its ability to reamplify on benzoate. As previously mentioned, when the *recA*⁺ strain carrying two copies was diluted and spread onto a benzoate medium agarose plate where single colonies were separated (unlike the original selective conditions), large size colonies grew within 4-7 days at 22°C. Furthermore, the more copies of the amplicon the *recA*⁺ strains started with the more quickly it formed large colonies on benzoate (data not shown). In contrast, when the cells from the *recA* mutant strain with two copies were diluted and spread-plated onto selective media, they both formed small pinpoint colonies after approximately 7 days incubation. However, these colonies always took several days longer to form visible tiny colonies and they never increased in size, as they did in the *recA*⁺ two copy strain. Cells from 10 colonies of each of these strains were transferred from colonies on different benzoate plates and grown in liquid succinate MM for DNA isolation to be analyzed by PFGE and Southern hybridization. This analysis revealed that the *recA*⁺ cells with multiple copies had regenerated

amplification on the benzoate plate, while the *recA* mutant cells with two copies had not regenerated amplification. Amplification by the *recA*⁺ strain was indicated by a ladder of fragments, rather than by a single intense band. In contrast, the *recA* mutant strain did not have this ladder, nor did its two copies deamplify detectably to single copy; however, the two copies of the amplicon appeared to remain static. Therefore, RecA appears necessary for the second step of amplification but not for the first.

Oxolinic acid and other DNA damaging agents increase the amplification mutation rate

As previously mentioned, amplification in the *Acinetobacter* system may be explained by a natural selective process that does not necessarily imply a mechanism that is inducible by stress caused by the selective medium. However, in other systems amplification appears to be inducible by agents independent of the selective media. Pretreatment of the cells with agents that induce double-stranded DNA breaks (DSBs) increase the frequency of mammalian gene amplification (Mondello *et al.*, 2002). In these mammalian cells, DSBs are mainly repaired through the non-homologous end-joining pathways (NHEJ). DSBs that are not repaired are believed to initiate gene amplification possibly by providing DNA ends to be directly available for recombination events that can lead to gene amplification (i.e. duplication formation). Alternatively, recent discoveries suggest that DSBs may activate gene amplification pathway(s) (Mondello *et al.*, 2002). In order to investigate whether amplification was inducible in the *Acinetobacter* system, the parent strain was grown in nonselective media containing different mutagenic agents and treatments prior to selection. The effect of these agents on amplification induction was interpreted by measuring the Ben⁺ mutation frequency.

As previously mentioned, the duplication junctions isolated from the *cat* gene amplification mutants resulted from illegitimate recombination that was site-specific in some cases. While it is currently unknown by which method such duplications form, one possibility is the involvement of DNA gyrase. DNA gyrase, a member of the (type II) topoisomerase class of enzymes, is an essential bacterial enzyme that catalyzes the introduction of negative superhelical turns into closed-circuit dsDNA. DNA gyrase relaxes supercoiled DNA by making a DSB, followed by strand passage, and completed by religation of the DSB. A DSB that is not correctly repaired may cause a duplication if there is recombination between two gyrase cut sites located on sister chromosomes. Previous studies of the effects of a DNA gyrase inhibitor, oxolinic acid, as well as a DNA gyrase mutation on deletion formation in *E. coli* indicated that DNA gyrase participates in illegitimate recombination *in vivo* (Miura-Masuda and Ikeda, 1990). DNA gyrase has also been shown to cut by site-specific illegitimate recombination. Gyrase activity has been shown to be inhibited by the DNA-damaging agent, oxolinic acid. This agent is believed to target the GyrA subunit of the DNA gyrase holoenzyme, inhibiting supercoiling by repressing the religation sequence-specific cuts on the DNA duplex. Thus, oxolinic acid induction of gene amplification was tested.

In cancer cell lines, cells can respond transiently to agents (e.g. mitomycin C) or treatments (e.g. UV light) that damage DNA or arrest DNA synthesis by increasing their rate of gene amplification as much as 10- to 100-fold (Stark and Wahl, 1984; Schimke, 1988). Mitomycin C (MMC), an antitumor agent isolated from *Streptomyces* cultures, is used in chemotherapy (Stark, 1993). It interacts with DNA by intercalation and adduct formation. These actions also trigger the SOS response in *E. coli*, resulting in the concerted induction of several DNA repair and recombination activities (Khil and Camerini-Otero, 2002). In ADP1,

mitomycin C induces expression of *recA*, while UV does not (Rauch *et al.*, 1996). Thus, we tested the ability of mitomycin C to induce gene amplification in *Acinetobacter*. Another DNA damaging agent shown to induce *recA* expression in *Acinetobacter*, methyl-methanesulfonate (MMS), was also tested.

As shown in figure 4.4, pregrowth of the parent strain in the presence of DNA-damaging agents yielded an increased Ben⁺ mutation frequency by up to two to three orders of magnitude. While oxolinic acid caused an increase in the mutation frequency on the average of 35-fold, mitomycin C and MMS caused an increase of approximately 206 and 264-fold respectively. UV treatment did not appear to affect the mutation frequency. This presumably reflects ADP1's UV resistance (Rauch *et al.*, 1996).

In an attempt to determine whether the initial duplication step of amplification was being induced by a specific DNA-damaging agent, PCR was used to estimate the frequency of formation for four of the site-specific duplication junctions in cells grown in succinate in the presence of these agents. Based on the four junctions tested, site specific duplication in unselected populations grown in the presence of DNA-damaging agents does not appear to increase the frequency of these duplication events (Reams and Neidle unpublished data). This suggests that the DNA damaging agents do not induce the duplication step, at least for these four junctions. Since these agents increase the mutation frequency but not the occurrence of duplications, we propose that the DNA damaging agents induce expression of the genes responsible for the second step of gene amplification. Evidence supporting this is the induction of *recA* expression by both mitomycin C and MMS, but not by UV (Rauch *et al.*, 1996).

To test further whether the DNA-damaging agents were directly affecting the duplication process, 14 junctions were isolated from oxolinic acid induced gene amplification mutants.

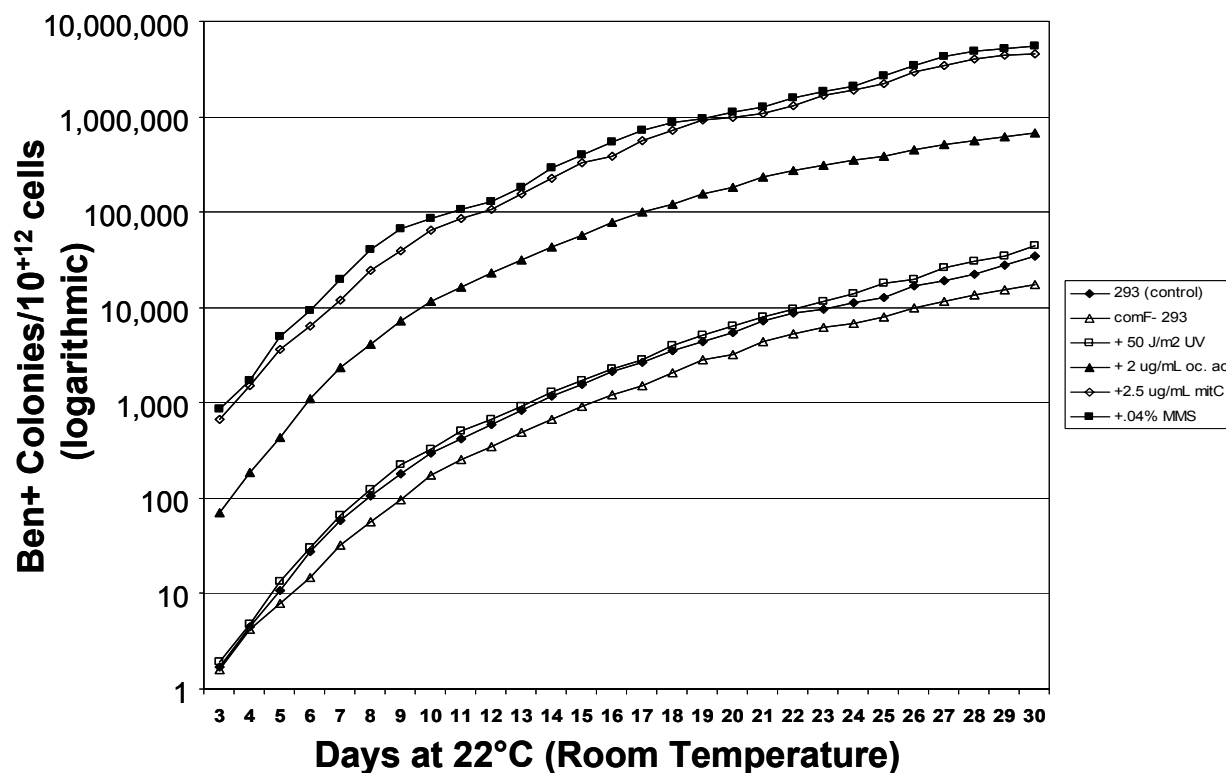


Fig. 4.4 DNA-damaging agents induce gene amplification.

The total number of visible Ben⁺ colonies, counted at the end of each day after inoculation, was compared to the viable cell count determined from a sample of the same culture grown on succinate medium. Each point represents the average value from at least 5 independently repeated trials with 10 to 20 plates per trial. The parent strain, ACN293, grown in the absence of mutations or DNA-damaging agents/treatments, was used as a control and is represented by 293. In treated cultures, the parent strain was pre-grown in succinate medium in the presence of one of the following DNA-damaging agents: 0.04% methyl-methanesulphonate (MMS), 2.5 µg/mL mitomycin C (mit C), or 2 µg/mL oxolinic acid (oc. ac.). UV treatment was issued as 50J/m². The *comF* mutant of the parent strain (*comF*⁻ 293) was pre-grown in succinate medium.

Sequence from these mutants revealed that all 14 recombination events were unique and had resulted from non-homologous recombination involving between 0-3 bp (Reams and Neidle unpublished data). No site-specificity was found. If oxolinic acid acted in the same way as it did in *E. coli*, that is as a gyrase inhibitor causing illegitimate recombination at gyrase cut sites, one may presume that oxolinic acid would directly induce duplications at gyrase cut sites. However, no site-specificity indicative of gyrase cut sites was found. On the other hand, it is still possible that these 14 junctions may reflect duplications initiated directly by oxolinic acid. More studies will need to be performed to characterize gyrase cut sites.

Natural transformability of ADP1 does not appear to be necessary for gene amplification

The natural transformation of *Acinetobacter* sp. strain ADP1 has facilitated the genetic manipulation of these strains. Indeed, a novel assay based on ADP1's natural transformability allowed for the isolation of 86 duplication junction sequences (Reams and Neidle, unpublished data). However, it is unknown whether this transformability contributes to the process of amplification. One possibility is that while cells are sitting on the selective plate, some cells are lysing, releasing their genomic DNA. DNA fragments carrying a duplication junction specific to the *cat* genes selection may transform competent cells and induce gene amplification. This induction is proposed to occur similarly to that in the transformation assay. To investigate this hypothesis, we examined whether a parent strain deficient in natural transformation was able to form amplification mutants under the original selection conditions at a similar rate to that of the parent strain. There are several genes involved in competency for ADP1. One gene, *comF*, results in a 1,000-fold decrease in transformability when knocked out (Busch *et al.*, 1999). Therefore, a *comF* mutant of the parent strain was constructed. The transformability of this

parent strain with a *comF* insertion was measured to be approximately 1,000-2,000-fold less transformable than the *comF*⁺ parent strain. When the *comF* mutant parent strain was put under the original selective conditions, the Ben⁺ mutation rate was similar to that of the *comF*⁺ parent strain (fig. 4.4). PFGE and Southern hybridization analysis on 18 Ben⁺ strains that lacked *comF* revealed that all of these strains carried *cat* gene amplification. Therefore, this indicates that transformability does not significantly contribute to the amplification process.

Proposed model for two-step gene amplification

Based on our results, Figure 4.5 depicts the process of two-step gene amplification of the *cat* genes in *Acinetobacter*. We believe several of the phenomena included in this model may be applicable to many other gene amplification systems. While some of the concepts in this figure have been previously proposed, until now, there has not been the experimental proof necessary to provide evidence for these concepts.

As shown in figure 4.5, the initial *cat* gene duplication occurs via illegitimate recombination which appears to be RecA-independent. While short homologies were found at some of these duplication recombination sites, most duplication events did not require homology. Site-specificity can be involved in these recombination events. The dependence of *cat* gene duplication formation on illegitimate recombination may reflect the restrictions of the selection where the upstream terminus of the amplicon must lie somewhere within an approximate 5.5 kb region between the *benA* and *catA* transcriptional start sites.

The generation of higher amplification is only apparent by detectable methods after selection. This second step of amplification is the cause for the delay in original colony formation under selective conditions and may therefore be considered the rate-limiting step. It

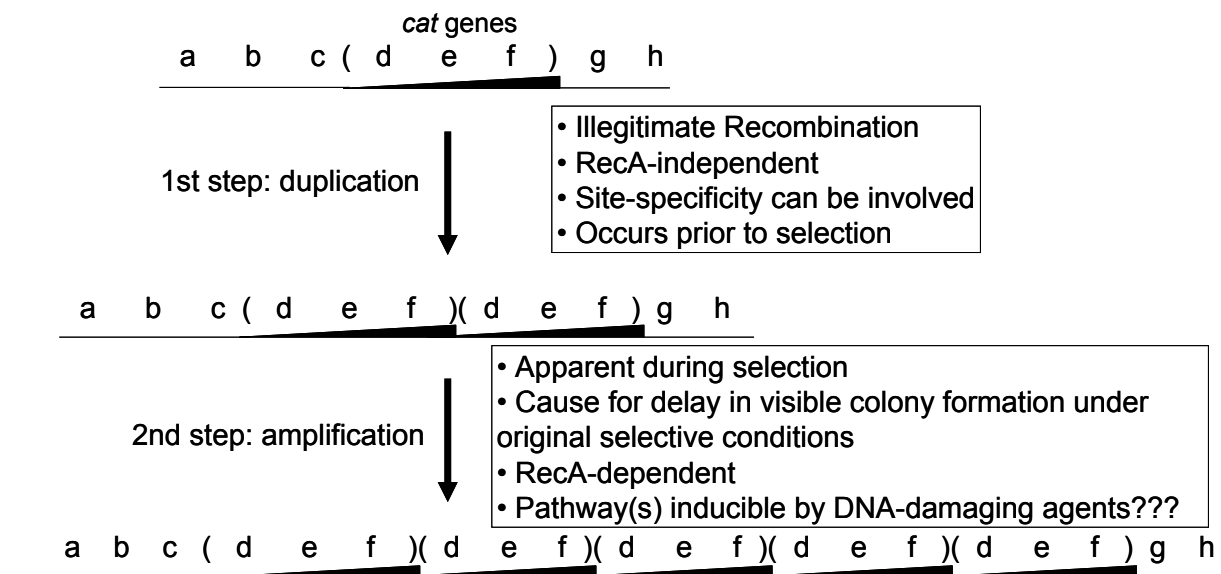


Fig. 4.5 Proposed model for two-step amplification in *cat* gene amplification of *Acinetobacter* sp. strain ACN293.

should be emphasized that the phrase “rate-limiting step” is a relative phrase in context only to the original selective conditions. For example, amplification can not occur until duplication has taken place. In fact, cells that do not carry a preexisting *cat* duplication can not undergo amplification when grown on selective plates. In contrast, when cells carrying a *cat* duplication are diluted and spread out on the benzoate plate, they will generate amplification within only a few days (≤ 7 days). This is much faster than the average amplification mutant colony under the original selective conditions. In this context, the initial duplication may be reversely considered the “rate-limiting step”.

The generation of higher amplification from a preexisting duplication is *recA*-dependent suggesting that it is a result of a series of sister strand exchanges. Conversely, it is still possible that *cat* gene amplification could occur via a rolling circle model (Petit *et al.*, 1992). In this fashion, a duplication would most likely first be generated. Following a double-stranded break at the downstream terminus of the second amplicon, RecA may allow for a homology hunt of the second amplicon to loop back with the first amplicon to prime the “do-loop” mechanism (Roth *et al.*, 1996). The alternative possibility, that the rolling circle mechanism is initiated prior to duplication formation, is highly unlikely due to lack of homology at the junction sites which would be the priming sites for the rolling circle. Also, the second steps dependence on RecA strongly suggests the involvement of homologous recombination, which this latter rolling circle mechanism presumable does not depend.

To what extent do DNA-damaging agents induce gene amplification? We believe that DNA-damaging agents increase the amplification rate by inducing expression of the pathway(s) for the generation of higher amplification from a preexisting duplication. This may be primarily based on the ability of these agents to induce *recA* expression. Clearly, more work is needed to

elucidate the role of DNA-damaging agents. This proposal may have some implications in how cancer is generated and treated. DNA-damaging agents, also called carcinogens, may cause cancer by inducing these amplification pathways, as opposed to causing the availability of DSB ends needed for duplication formation. Furthermore, in cancer treatment, chemotherapy and radiation cause a high amount of DNA-damage to both cancer and non-cancerous cells. Such treatment may also induce the pathways for the second step of amplification. Since the same pathways are believed to be responsible for amplification as well as deamplification, these treatments may be relatively curative because they cause deamplification by increasing gene expression. Much work needs to be done to elucidate how these pathways are regulated. With this future research, hopefully we can find compounds/treatments more efficient at inducing deamplification pathway(s) and less toxic to non-cancerous cells than chemotherapy and radiation.

Experimental Procedures

Bacterial strains, plasmids, and growth condition, and DNA manipulations

Acinetobacter spontaneous amplification mutants were derived from ACN293. These mutants arose on minimum media plates with 2 mM benzoate as the carbon source while incubated at 22°C. *E. coli* DH5 α and Top10 (Invitrogen) were used as plasmid hosts (Grant *et al.*, 1990). Unless otherwise noted, bacteria were cultured at 37°C in Luria-Bertani medium or minimal medium with 10 mM succinate or 2 mM benzoate as the carbon source (Sambrook *et al.*, 1989; Shanley *et al.*, 1986). Antibiotics were added as needed at the following final concentrations: kanamycin, 40 μ g/ml; streptomycin, 13 μ g/ml; spectinomycin, 13 μ g/ml. Cells grown in the presence of DNA damaging agents were first grown in a 10 mL overnight culture of succinate

minimal media without drugs. From this culture, 1 mL was transferred into 125 mL of fresh succinate minimal media with or without DNA-damaging agents. DNA-damaging agents were added at the following final concentrations: oxolinic acid, 2 µg/mL; mitomycin C, 2.5 µg/mL; methyl-methanesulfonate, 0.04%. These cells were grown overnight and plated onto benzoate minimal media plates. Cells that underwent UV treatment were exposed at a dose of 50 J/m² in open Petri dishes immediately upon plating onto benzoate plates.

Plasmid and library construction

To make plasmid libraries, *Acinetobacter* chromosomal DNA was purified with an Easy-DNA kit (Invitrogen). Samples of DNA from each mutant were digested with either *EcoRI* or *XbaI*, ligated to a similarly digested sample of the pZero-2.2 cloning vector (Invitrogen) and used to transform *E. coli*. This ColE1-type vector, which is lethal in *E. coli* without the insertion of cloned DNA, is designed to select kanamycin-resistant host cells that all carry recombinant plasmids.

Selection of oxolinic acid induced ACN293-derived Ben⁺ mutants and transformation assay to detect DNA that confers growth on benzoate to ACN293

A previously described transformation assay was used to screen libraries to locate the junction sequences from amplification isolates (Reams and Neidle, 2003).

Preparation of intact genomic DNA and restriction digestion for TAFE analysis

Plugs were prepared as described in Gralton *et al.* (1997).

Pulsed field gel electrophoresis and Southern hybridization methods

DNA samples were prepared and electrophoretically separated by the pulsed-field method of transverse alternating field electrophoresis (TAFE, Geneline II, Beckman) as previously described (Gralton *et al.*, 1997). Southern hybridization methods were carried out as indicated by Gralton *et al.*, 1997. Nonradioactive probes, prepared with digoxigenin and a random-primed labeling system (Genius Systems, Boehringer Mannheim), were hybridized to target sequences, and detected with anti-digoxigenin alkaline phosphatase conjugates and chemiluminescent substrates. Hybridization signal intensities on exposed X-ray films were digitized and calibrated on a UVP Lab Products gel documentation system according to the directions of the LabWorks image acquisition and analysis software.

Measurement of amplicon length

For all amplification mutants, amplicon length was estimated by PFGE and Southern hybridization as previously described (Reams and Neidle, unpublished results, please see chapter 3). For the 87 amplification mutants whose junctions were isolated, the amplicon sizes were verified by genome sequence analysis. For these mutants, the junction endpoints were sequenced over and the distance between them was calculated based on genome sequence (www.genoscope.com). The 446 bp deletion in *catM* was taken into account in this measurement.

PCR assays

Concentrated chromosomal DNA from the original ACN293 culture that ACN was derived from, along with other independent populations of ACN293 and ADP1 were prepared using phenol

extractions. Concentrated chromosomal DNA was digested with 1 unit *EcoRI* in a 10 μ L total volume reaction. Digests were incubated for 30 minutes at 37°C. After 30 minutes, an additional 0.5 unit of *EcoRI* was added to each digest and incubated for an additional 30 minutes at 37°C. Upon digest completion the restriction enzymes were heat-inactivated at 68°C for 15 minutes and diluted with 90 μ L of sterilized ddH₂O. DNA concentrations were determined using a fluorimeter assay. DNA was standardized to 30 ng/10 μ L and subject to a series of 10-fold dilutions. PCR reactions were performed in a 50 μ L total volume. In addition to template DNA, each PCR reactions contained the following reagents: 0.5 μ M of each primer, 100 μ M of each dNTP, 1X Faststart Taq buffer, and 1 unit Faststart Taq polymerase. The PCR reactions were performed in mini-epindorf tubes on a MJ Research MiniCycler with a hot bonnet. The PCR reactions were initiated by heating to 95°C for 4 minutes. The remainder of the PCR was conducted for 40 cycles at 94°C for 45 seconds (denaturation), 55°C for 45 seconds (annealing), 72°C for 30 seconds (elongation) followed by a 7 minute incubation at 72°C to complete the synthesis. The following primers were used in the PCR to assay the initial recombination event: 7255up20 5' GCAAACAACAGGCCAAGC 3' and 7255down277 5' CAATATCTAAAAAACCGAGG 3'.

Reconstruction Experiments

Reconstruction experiments were performed by plating a known number (approximately 100) of derivative ACN277 strains carrying different amplicon copy numbers (15, 8, 3, or 2) with about 5×10^9 ACN293 (scavengers) cells as in the original amplification selection. Before plating, each population of cells was grown in 10 mM succinate minimal medium. Background amplification rate was determined by plating scavengers without any ACN277 derivative cells.

Viable counts were determined from diluted cultures spread plated onto succinate MM plates. In each set of experiments, the same overnight culture of scavengers was plated with the derivative strains. Each set of experiments was repeated seven times.

CHAPTER 5

CONCLUSION

During my studies of the soil bacterium *Acinetobacter* sp. strain ADP1, I discovered a selection for gene amplification mutants. These studies provide the first and only system for studying gene amplification in *Acinetobacter*. In the past, our lab has focused specifically on ADP1 due to its ability to degrade aromatic compounds such as benzoate. Since these compounds are similar in structure to pollutants, these strains have potential applications in bioremediation.

Gene amplification has generated considerable interest in recent years due to the discovery that the amplification of cellular oncogenes within cancer cells contributes in important ways to tumorigenesis and metastasis. Also, the importance of gene amplification in pathogenic bacteria is underscored by its correlation with increased drug resistance and virulence. Although gene amplification has been observed in a wide variety of hosts ranging from viruses to humans, the underlying mechanism(s) remain unclear for any organism.

In the past, gene amplification has been a difficult process to study, due primarily to the lack of a good system to study this process. For several reasons, the mutant selection in *Acinetobacter* described in this thesis is an ideal system for studying gene amplification. First, one of the great advantages of working with ADP1 is its highly efficient natural transformation system. This competency facilitates the designing of mutants by allowing these strains to readily uptake free DNA in the environment. Once inside the cell these DNA segments can be integrated into the chromosome, usually by homologous recombination. Thus, mutants can be

constructed by designing segments of DNA and simply adding them to a culture of the recipient strain. Successful transformation may then be selected or screened. Moreover, gene amplification can be induced by transforming the parent strain with a junction fragment containing both ends of an amplicon. Using this trait, a transformation assay was designed that takes advantage of the natural competence of these *Acinetobacter* strains so that the precise endpoints of the amplicons could be determined.

Another advantage in studying ADP1 is the availability of its genome sequence. Recently, two companies have independently sequenced the approximate 3.6 Mb chromosome of ADP1: Genoscope (France) and Integrated Genomics (Chicago, IL). Genoscope has made their ADP1 sequence available to the public on the internet without charge (www.genoscope.com). While Integrated Genomics charges a fee for accessing their sequence, they provide an annotated version of the genome. Also, this company offers a much more advanced computer program allowing more sophisticated sequence analysis of the ADP1 genome, along with several other sequenced genomes (www.integratedgenomics.com).

In addition to being naturally transformable and having available genome sequence, ADP1 is non-pathogenic and genetically tractable making these strains relatively safe and easy to utilize. The fine precision of a bacterial genetic system enables this system's use as a paradigm for gene amplification. Furthermore, the selection described in this thesis is virtually exclusive for mutants carrying gene amplification. This trait facilitates the analysis of engineered mutations on their direct effects to the gene amplification process. For example, the removal of RecA completely abolishes the yield of mutants in this selection. In this manner, RecA was demonstrated to be essential for gene amplification. Moreover, RecA was shown to be required for the second step of gene amplification and not for the first step. This *Acinetobacter* system

may provide the means for testing the involvement of other recombinational repair genes in amplification. Pursuing the identification of pathways involved in amplification and deamplification may transform this model system into an impetus for the pursuit of identifying these pathways in higher cells.

This gene amplification system also has evolutionary significance. For example, this system demonstrates that the co-amplification of clustered functionally-related genes provides a selective advantage under certain environments. These findings presented new evidence that supports the “co-amplification theory”, which states that the co-amplification of functionally related genes is a driving force in the evolution of higher-order gene clusters. In the future, the significance of this theory and the contribution of gene amplification in the evolution of gene clusters certainly will be investigated further.

Another evolutionary theory explored in this thesis is the controversial theory of “adaptive amplification”. According to this theory, amplification is actually induced by stress created by the selective medium. Controversy has surrounded this theory, due to its implications towards the role of the environment in causing genetic variation during evolution. My results indicated that amplification could be explained by a natural selection process. Thus, these mutants did not arise adaptively.

In summary, these gene amplification mutants provide a new model system for studying a process with important implications to evolution, drug resistance, increased virulence, bioremediation, and cancer. This model system may provide valuable insights for understanding aspects of pathogenesis, as well as for understanding gene amplification in higher cells. Treatment for a variety of human diseases and cancers should improve significantly from a better understanding of this important process.

REFERENCES

- Acheson, D.W., Reidl, J., Zhang, X., Keusch, G.T., Mekalanos, J.J., and Waldor, M.K. (1998)** *In vivo* transduction with shiga toxin 1-encoding phage. *Infect Immun* **66**: 4496-4498.
- Albertini, A.M., Hofer, M., Calos, M.P., and Miller, J.H. (1982)** On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* **29**: 319-328.
- Altenbuchner, J., and Cullum, J. (1987)** Amplification of cloned genes in *Streptomyces*. *Bio/technology* **5**: 1328-1329.
- Anderson, R.P., and Roth, J.R. (1981)** Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (*rrn*) cistrons. *Proc Natl Acad Sci U S A* **78**: 3113-3117.
- Anderson, R.P., and Roth, J.R. (1977)** Tandem genetic duplications in phage and bacteria. *Annu Rev Microbiol* **31**: 473-505.
- Andersson, D.I., Slechta, E.S., and Roth, J.R. (1998)** Evidence that gene amplification underlies adaptive mutability of the bacterial *lac* operon. *Science* **282**: 1133-1135.
- Assinder, S.J., and Williams, P.A. (1990)** The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv Microb Physiol* **31**: 1-69.
- Betley, M.J., and Mekalanos, J.J. (1985)** Staphylococcal enterotoxin A is encoded by phage. *Science* **229**: 185-187.
- Bradford, M.M. (1976)** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.
- Brown, T.A. (2002)** *Genomes 2*. Bios Scientific Publishers Limited.
- Brzostowicz, P.C., Reams, A.B., Clark, T.J., and Neidle, E.L. (2003)** Transcriptional cross-regulation of the catechol and protocatechuate branches of the beta-ketoadipate pathway contributes to carbon source-dependent expression of the *Acinetobacter* sp. strain ADP1 *pobA* gene. *Appl Environ Microbiol* **69**: 1598-1606.
- Buchan, A., Neidle, E.L. and Moran, M.A. (2003)** Dynamic organization of genes of the beta-ketoadipate pathway in members of the marine *Roseobacter* Lineage. *In Press*.

- Bundy, B.M., Campbell, A.L., and Neidle, E.L. (1998)** Similarities between the *antABC*-encoded anthranilate dioxygenase and the *benABC*-encoded benzoate dioxygenase of *Acinetobacter* sp. strain ADP1. *J Bacteriol* **180**: 4466-4474.
- Bundy, B.M., Collier, L.S., Hoover, T.R., and Neidle, E.L. (2002)** Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proc Natl Acad Sci U S A* **99**: 7693-7698.
- Busch, S., Rosenplanter, C., and Averhoff, B. (1999)** Identification and characterization of ComE and ComF, two novel pilin-like competence factors involved in natural transformation of *Acinetobacter* sp. strain BD413. *Appl Environ Microbiol* **65**: 4568-4574.
- Clark, T.J., Momany, C., and Neidle, E.L. (2002)** The *benPK* operon, proposed to play a role in transport, is part of a regulon for benzoate catabolism in *Acinetobacter* sp. strain ADP1. *Microbiology* **148**: 1213-1223.
- Collier, L.S., Gaines, G.L., 3rd, and Neidle, E.L. (1998)** Regulation of benzoate degradation in *Acinetobacter* sp. strain ADP1 by BenM, a LysR-type transcriptional activator. *J Bacteriol* **180**: 2493-2501.
- Corn, P.G., Anders, J., Takala, A.K., Kayhty, H., and Hoiseth, S.K. (1993)** Genes involved in *Haemophilus influenzae* type b capsule expression are frequently amplified. *J Infect Dis* **167**: 356-364.
- Cornet, F., Mortier, I., Patte, J., and Louarn, J.M. (1994)** Plasmid pSC101 harbors a recombination site, *psi*, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site *dif*. *J Bacteriol* **176**: 3188-3195.
- Cosper, N.J., Collier, L.S., Clark, T.J., Scott, R.A., and Neidle, E.L. (2000)** Mutations in *catB*, the gene encoding muconate cycloisomerase, activate transcription of the distal *ben* genes and contribute to a complex regulatory circuit in *Acinetobacter* sp. strain ADP1. *J Bacteriol* **182**: 7044-7052.
- Cotter, P.A., DiRita, V.J. (2000)** Bacterial virulence gene regulation: An evolutionary perspective. *Annu Rev Microbiol* **54**: 519-565.
- Davis, B.M., and Waldor, M.K. (2003)** Filamentous phages linked to virulence of *Vibrio cholerae*. *Curr Opin Microbiol* **6**: 35-42.
- de la Cruz, F., and Davies, J. (2000)** Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol* **8**: 128-133.
- de Vries, J., and Wackernagel, W. (2002)** Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc Natl Acad Sci U S A* **99**: 2094-2099.

- Doolittle, R.F. (1987)** The evolution of the vertebrate plasma proteins. *Biol Bull* **172**:269-283.
- Dorai, H., and Moore, G.P. (1987)** The effect of dihydrofolate reductase-mediated gene amplification on the expression of transfected immunoglobulin genes. *J Immunol* **139**: 4232-4241.
- Edlund, T., and Normark, S. (1981)** Recombination between short DNA homologies causes tandem duplication. *Nature* **292**: 269-271.
- Ehrlich, S.D. (1989)** Illegitimate recombination in bacteria. In *Mobile DNA*. Berg, D.E. and Howe, M.E. (eds). Washington, DC: ASM Press, pp. 799-832.
- Ehrlich, S.D., Bierne, H., d'Alencon, E., Vilette, D., Petranovic, M., Noirot, P., and Michel, B. (1993)** Mechanisms of illegitimate recombination. *Gene* **135**: 161-166.
- Eklund, M.W., and Poysky, F.T. (1974)** Interconversion of type C and D strains of *Clostridium botulinum* by specific bacteriophages. *Appl Microbiol* **27**: 251-258.
- Eraso, J.M., and Kaplan, S. (1994)** *prpA*, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*. *J Bacteriol* **176**: 32-43.
- Finnerty, J.R., and Martindale, M.Q. (1999)** Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. *Evol Dev* **1**: 16-23.
- Flores, M., Mavingui, P., Perret, X., Broughton, W.J., Romero, D., Hernandez, G., Davila, G., and Palacios, R. (2000)** Prediction, identification, and artificial selection of DNA rearrangements in *Rhizobium*: toward a natural genomic design. *Proc Natl Acad Sci U S A* **97**: 9138-9143.
- Fornace, A.J., Jr., Alamo, I., Jr., and Hollander, M.C. (1988)** DNA damage-inducible transcripts in mammalian cells. *Proc Natl Acad Sci U S A* **85**: 8800-8804.
- Foster, P.L. (1999)** Mechanisms of stationary phase mutation: a decade of adaptive mutation. *Annu Rev Genet* **33**: 57-88.
- Franklin, N.C. (1967)** Extraordinary recombinational events in *Escherichia coli*. Their independence of the *rec+* function. *Genetics* **55**: 699-707.
- Freeman, V.J. (1951)** Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *J Bacteriol* **162**: 668-675.
- Gaines, G.L., 3rd, Smith, L., and Neidle, E.L. (1996)** Novel nuclear magnetic resonance spectroscopy methods demonstrate preferential carbon source utilization by *Acinetobacter calcoaceticus*. *J Bacteriol* **178**: 6833-6841.

- Galitski, T., and Roth, J.R. (1995)** Evidence that F plasmid transfer replication underlies apparent adaptive mutation. *Science* **268**: 421-423.
- Gerischer, U., D'Argenio, D.A., and Ornston, L.N. (1996)** IS1236, a newly discovered member of the IS3 family, exhibits varied patterns of insertion into the *Acinetobacter calcoaceticus* chromosome. *Microbiology* **142** (Pt 7): 1825-1831.
- Ghosal, D., and You, I.S. (1988)** Gene duplication in haloaromatic degradative plasmids pJP4 and pJP2. *Can J Microbiol* **34**: 709-715.
- Goldberg, I., and Mekalanos, J.J. (1986)** Effect of a *recA* mutation on cholera toxin gene amplification and deletion events. *J Bacteriol* **165**: 723-731.
- Gralton, E.M., Campbell, A.L., and Neidle, E.L. (1997)** Directed introduction of DNA cleavage sites to produce a high-resolution genetic and physical map of the *Acinetobacter* sp. strain ADP1 (BD413UE) chromosome. *Microbiology* **143** (Pt 4): 1345-1357.
- Grant, S.G., Jessee, J., Bloom, F.R., and Hanahan, D. (1990)** Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A* **87**: 4645-4649.
- Gregg-Jolly, L.A., and Ornston, L.N. (1990)** Recovery of DNA from the *Acinetobacter calcoaceticus* chromosome by gap repair. *J Bacteriol* **172**: 6169-6172.
- Gregg-Jolly, L.A., and Ornston, L.N. (1994)** Properties of *Acinetobacter calcoaceticus recA* and its contribution to intracellular gene conversion. *Mol Microbiol* **12**: 985-992.
- Grondin, K., Haimeur, A., Mukhopadhyay, R., Rosen, B.P., and Ouellette, M. (1997)** Co-amplification of the gamma-glutamylcysteine synthetase gene *gshI* and of the ABC transporter gene *pgpA* in arsenite-resistant *Leishmania tarentolae*. *EMBO J* **16**: 3057-3065.
- Guttersen, N.I., and Koshland, D.E., Jr. (1983)** Replacement and amplification of bacterial genes with sequences altered *in vitro*. *Proc Natl Acad Sci U S A* **80**: 4894-4898.
- Gyles, C., So, M., and Falkow, S. (1974)** The enterotoxin plasmids of *Escherichia coli*. *J Infect Dis* **130**: 40-49.
- Hacker, J., and Kaper, J.B. (2000)** Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**: 641-679.
- Hagblom, P., Segal, E., Billyard, E., and So, M. (1985)** Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. *Nature* **315**: 156-158.

- Harms, C.T., Armour, S.L., DiMaio, J.J., Middlesteadt, L.A., Murray, D., Negrotto, D.V., Thompson-Taylor, H., Weymann, K., Montoya, A.L., Shillito, R.D., and *et al.* (1992)** Herbicide resistance due to amplification of a mutant acetohydroxyacid synthase gene. *Mol Gen Genet* **233**: 427-435.
- Hartman, P.E., Loper, J.C., and Serman, D. (1960)** Fine structure mapping by complete transduction between histidine-requiring *Salmonella* mutants. *J Gen Microbiol* **22**: 323-353.
- Harwood, C.S., and Parales, R.E. (1996)** The beta-ketoadipate pathway and the biology of self-identity. *Annu Rev Microbiol* **50**: 553-590.
- Hastings, P.J., Bull, H.J., Klump, J.R., and Rosenberg, S.M. (2000)** Adaptive amplification: an inducible chromosomal instability mechanism. *Cell* **103**: 723-731.
- Heath, J.D., and Weinstock, G.M. (1991)** Tandem duplications of the *lac* region of the *Escherichia coli* chromosome. *Biochimie* **73**: 343-352.
- Hendricks, M.B., Luchette, C.A., and Banker, M.J. (1989)** Enhanced expression of an immunoglobulin-based vector in myeloma cells mediated by coamplification with a mutant dihydrofolate reductase gene. *Bio/technology* **7**: 1271-1274.
- Hendrickson, H., Slechts, E.S., Bergthorsson, U., Andersson, D.I., and Roth, J.R. (2002)** Amplification-mutagenesis: evidence that "directed" adaptive mutation and general hypermutability result from growth with a selected gene amplification. *Proc Natl Acad Sci U S A* **99**: 2164-2169.
- Hoiseth, S.K., Moxon, E.R., and Silver, R.P. (1986)** Genes involved in *Haemophilus influenzae* type b capsule expression are part of an 18-kilobase tandem duplication. *Proc Natl Acad Sci U S A* **83**: 1106-1110.
- Ikeda, H., Aoki, K., and Naito, A. (1982)** Illegitimate recombination mediated *in vitro* by DNA gyrase of *Escherichia coli*: structure of recombinant DNA molecules. *Proc Natl Acad Sci U S A* **79**: 3724-3728.
- Inselburg, J. (1967)** Formation of deletion mutations in recombination-deficient mutants of *Escherichia coli*. *J Bacteriol* **94**: 1266-1267.
- Jacob, F., Perrin, D., Sanchez, C., and Monod, J. (1960)** Operon: a group of genes with the expression coordinated by an operator. *C R Hebd Seances Acad Sci* **250**: 1727-1729.
- Jacob, F., and Monod J. (1962)** On the regulation of gene activity. *Cold Spring Harbor Symp. Quant. Bio.* **26**: 193-211.

- Jones, J.D., Weller, S.C., and Goldsbrough, P.B. (1994)** Selection for kanamycin resistance in transformed petunia cells leads to the co-amplification of a linked gene. *Plant Mol Biol* **24**: 505-514.
- Jones, R.M., Collier, L.S., Neidle, E.L., and Williams, P.A. (1999)** *areABC* genes determine the catabolism of aryl esters in *Acinetobacter* sp. Strain ADP1. *J Bacteriol* **181**: 4568-4575.
- Jones, R.M., Pagmantidis, V., and Williams, P.A. (2000)** *sal* genes determining the catabolism of salicylate esters are part of a supraoperonic cluster of catabolic genes in *Acinetobacter* sp. strain ADP1. *J Bacteriol* **182**: 2018-2025.
- Jordan, I.K., Makarova, K.S., Spouge, J.L., Wolf, Y.I., and Koonin, E.V. (2001)** Lineage-specific gene expansions in bacterial and archaeal genomes. *Genome Res* **11**: 555-565.
- Juni, E., and Janik, A. (1969)** Transformation of *Acinetobacter calcoaceticus* (*Bacterium anitratum*). *J Bacteriol* **98**: 281-288.
- Juni, E. (1978)** Genetics and physiology of *Acinetobacter*. *Annu Rev Microbiol* **32**: 349-371.
- Khil, P.P., and Camerini-Otero, R.D. (2002)** Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol Microbiol* **44**: 89-105.
- Klemba, M., Jakobs, B., Wittich, R.M., and Pieper, D. (2000)** Chromosomal integration of *tcb* chlorocatechol degradation pathway genes as a means of expanding the growth substrate range of bacteria to include haloaromatics. *Appl Environ Microbiol* **66**: 3255-3261.
- Kofoed, E., Bergthorsson, U., Slechta, E.S., and Roth, J.R. (2003)** Formation of an F' plasmid by recombination between imperfectly repeated chromosomal Rep sequences: a closer look at an old friend (F'(128) pro *lac*). *J Bacteriol* **185**: 660-663.
- Kok, R.G., Young, D.M., and Ornston, L.N. (1999)** Phenotypic expression of PCR-generated random mutations in a *Pseudomonas putida* gene after its introduction into an *Acinetobacter* chromosome by natural transformation. *Appl Environ Microbiol* **65**: 1675-1680.
- Kondrashov, F.A., Rogozin, I.B., Wolf, Y.I., and Koonin, E.V. (2002)** Selection in the evolution of gene duplications. *Genome Biol* **3**: RESEARCH0008.
- Kowalczykowski, S.C. (2000)** Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci* **25**: 156-165.
- Kroll, J.S., Loynds, B.M., and Moxon, E.R. (1991)** The *Haemophilus influenzae* capsulation gene cluster: a compound transposon. *Mol Microbiol* **5**: 1549-1560.

- Lai, C.Y., Baumann, L., and Baumann, P. (1994)** Amplification of *trpEG*: adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proc Natl Acad Sci U S A* **91**: 3819-3823.
- Lathe, W.C., 3rd, Snel, B., and Bork, P. (2000)** Gene context conservation of a higher order than operons. *Trends Biochem Sci* **25**: 474-479.
- Lawrence, J.G., and Roth, J.R. (1996)** Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**: 1843-1860.
- Lawrence, J.G. (2003)** Gene organization: selection, selfishness, and serendipity. *Annu Rev Microbiol* **57**: 419-440.
- Lee, J.M., and Sonnhammer, E.L. (2003)** Genomic gene clustering analysis of pathways in eukaryotes. *Genome Res* **13**: 875-882.
- Lin, R.J., Capage, M., and Hill, C.W. (1984)** A repetitive DNA sequence, *rhs*, responsible for duplications within the *Escherichia coli* K-12 chromosome. *J Mol Biol* **177**: 1-18.
- Lupski, J.R., Roth, J.R., and Weinstock, G.M. (1996)** Chromosomal duplications in bacteria, fruit flies, and humans. *Am J Hum Genet* **58**: 21-27.
- McDonough, M.A., and Butterton, J.R. (1999)** Spontaneous tandem amplification and deletion of the shiga toxin operon in *Shigella dysenteriae* 1. *Mol Microbiol* **34**: 1058-1069.
- Meagher, R.B., Ngai, K.L., and Ornston, L.N. (1990)** Muconate cycloisomerase. *Methods Enzymol* **188**: 126-130.
- Mekalanos, J.J. (1983)** Duplication and amplification of toxin genes in *Vibrio cholerae*. *Cell* **35**: 253-263.
- Mikesell, P., Ivins, B.E., Ristroph, J.D., and Dreier, T.M. (1983)** Evidence for plasmid-mediated toxin production in *Bacillus anthracis*. *Infect Immun* **39**: 371-376.
- Miura-Masuda, A., and Ikeda, H. (1990)** The DNA gyrase of *Escherichia coli* participates in the formation of a spontaneous deletion by *recA*-independent recombination *in vivo*. *Mol Gen Genet* **220**: 345-352.
- Mondello, C., Guasconi, V., Giulotto, E., and Nuzzo, F. (2002)** Gamma-ray and hydrogen peroxide induction of gene amplification in hamster cells deficient in DNA double strand break repair. *DNA Repair (Amst)* **1**: 483-493.
- Muller, T.A., Werlen, C., Spain, J., and Van Der Meer, J.R. (2003)** Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in *Ralstonia*. *Environ Microbiol* **5**: 163-173.

- Nash, H.A. (1996)** Site-specific recombination: integration, excision, resolution and inversion of defined DNA segments. In *Escherichia coli and Salmonella Cellular and Molecular Biology*. Neidhardt, F.C., Curtiss III, R., Ingraham, J.L., Lin, E.C.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M. and Umberger, H.E. (eds). Washington, DC: ASM Press, pp. 2363-2376.
- Neidle, E.L., Hartnett, C., and Ornston, L.N. (1989)** Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. *J Bacteriol* **171**: 5410-5421.
- Ngai, K.L., Neidle, E.L., and Ornston, L.N. (1990)** Catechol and chlorocatechol 1,2-dioxygenases. *Methods Enzymol* **188**: 122-126.
- O'Brien, A.D., Newland, J.W., Miller, S.F., Holmes, R.K., Smith, H.W., and Formal, S.B. (1984)** Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**: 694-696.
- Ogawa, N., and Miyashita, K. (1995)** Recombination of a 3-chlorobenzoate catabolic plasmid from *Alcaligenes eutrophus* NH9 mediated by direct repeat elements. *Appl Environ Microbiol* **61**: 3788-3795.
- Ogawa, N., and Miyashita, K. (1999)** The chlorocatechol-catabolic transposon Tn5707 of *Alcaligenes eutrophus* NH9, carrying a gene cluster highly homologous to that in the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. strain P51, confers the ability to grow on 3-chlorobenzoate. *Appl Environ Microbiol* **65**: 724-731.
- Ohno, S. (1970)** *Evolution by Gene Duplication*. Heidelberg: Springer Verlag.
- Parke, D., Garcia, M.A., and Ornston, L.N. (2001)** Cloning and genetic characterization of dca genes required for beta-oxidation of straight-chain dicarboxylic acids in *Acinetobacter* sp. strain ADP1. *Appl Environ Microbiol* **67**: 4817-4827.
- Parke, D., and Ornston, L.N. (2003)** Hydroxycinnamate (*hca*) catabolic genes from *Acinetobacter* sp. strain ADP1 are repressed by HcaR and are induced by hydroxycinnamoyl-coenzyme A thioesters. *Appl Environ Microbiol* **69**: 5398-5409.
- Peterson, B.C., and Rownd, R.H. (1983)** Homologous sequences other than insertion elements can serve as recombination sites in plasmid drug resistance gene amplification. *J Bacteriol* **156**: 177-185.
- Petit, M.A., Mesas, J.M., Noirot, P., Morel-Deville, F., and Ehrlich, S.D. (1992)** Induction of DNA amplification in the *Bacillus subtilis* chromosome. *EMBO J* **11**: 1317-1326.
- Pittard, J. (1996)** The various strategies within the TyrR regulation of *Escherichia coli* to modulate gene expression. *Genes Cells* **1**: 717-725.

- Pollard, S.L., and Holland, P.W. (2000)** Evidence for 14 homeobox gene clusters in human genome ancestry. *Curr Biol* **10**: 1059-1062.
- Prozorov, A.A. (2001)** Recombinant rearrangements of bacterial genome and adaptation to the environment. *Mikrobiologiya* **70**: 581-594.
- Rangnekar, V.M. (1988)** Variation in the ability of *Pseudomonas* sp. strain B13 cultures to utilize meta-chlorobenzoate is associated with tandem amplification and deamplification of DNA. *J Bacteriol* **170**: 1907-1912.
- Rauch, P.J., Palmen, R., Burds, A.A., Gregg-Jolly, L.A., van der Zee, J.R., and Hellingwerf, K.J. (1996)** The expression of the *Acinetobacter calcoaceticus* *recA* gene increases in response to DNA damage independently of RecA and of development of competence for natural transformation. *Microbiology* **142** (Pt 4): 1025-1032.
- Ravatn, R., Studer, S., Springael, D., Zehnder, A.J., and van der Meer, J.R. (1998)** Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. Strain B13. *J Bacteriol* **180**: 4360-4369.
- Reams, A.B., and Neidle, E.L. (2003)** Genome plasticity in *Acinetobacter*: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. *Mol Microbiol* **47**: 1291-1304.
- Ritchie, J.M., Wagner, P.L., Acheson, D.W., and Waldor, M.K. (2003)** Comparison of Shiga toxin production by hemolytic-uremic syndrome-associated and bovine-associated Shiga toxin-producing *Escherichia coli* isolates. *Appl Environ Microbiol* **69**: 1059-1066.
- Romero, D., and Palacios, R. (1997)** Gene amplification and genomic plasticity in prokaryotes. *Annu Rev Genet* **31**: 91-111.
- Romero-Arroyo, C.E., Schell, M.A., Gaines, G.L., 3rd, and Neidle, E.L. (1995)** *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. *J Bacteriol* **177**: 5891-5898.
- Rosenberg, S.L., and Hegeman, G.D. (1969)** Clustering of functionally related genes in *Pseudomonas aeruginosa*. *J Bacteriol* **99**: 353-355.
- Rosenberg, S.M., and Hastings, P.J. (2003)** Microbiology and evolution: modulating mutation rates in the wild. *Science* **300**: 1382-1383.
- Roth, J.R., Benson, N., Galitski, T., Haack, K., Lawrence, J.G., and Miesel, L. (1996)** Rearrangements of the bacterial chromosome: formation and applications. In *Escherichia coli and Salmonella Cellular and Molecular Biology*. Neidhardt, F.C., Curtiss III, R., Ingraham, J.L., Lin, E.C.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M. and Umberger, H.E. (eds). Washington, DC: ASM Press, pp. 2256-2276.

- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989)** *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press.
- Savelyeva, L., and Schwab, M. (2001)** Amplification of oncogenes revisited: from expression profiling to clinical application. *Cancer Lett* **167**: 115-123.
- Schimke, R.T. (1988)** Gene amplification in cultured cells. *J Biol Chem* **263**: 5989-5992.
- Segovia, M. (1994)** *Leishmania* gene amplification: a mechanism of drug resistance. *Ann Trop Med Parasitol* **88**: 123-130.
- Seoane, A., Sanchez, E., and Garcia-Lobo, J.M. (2003)** Tandem amplification of a 28-kilobase region from the *Yersinia enterocolitica* chromosome containing the *blaA* gene. *Antimicrobial Agents and Chemotherapy* **47**: 682-688.
- Shanley, M.S., Neidle, E.L., Parales, R.E., and Ornston, L.N. (1986)** Cloning and expression of *Acinetobacter calcoaceticus* *catBCDE* genes in *Pseudomonas putida* and *Escherichia coli*. *J Bacteriol* **165**: 557-563.
- Shimizu, H., Yamaguchi, H., Ashizawa, Y., Kohno, Y., Asami, M., Kato, J., and Ikeda, H. (1997)** Short-homology-independent illegitimate recombination in *Escherichia coli*: distinct mechanism from short-homology-dependent illegitimate recombination. *J Mol Biol* **266**: 297-305.
- Shyr, Y.Y., Hepburn, A.G., and Widholm, J.M. (1992)** Glyphosate selected amplification of the 5-enolpyruvylshikimate-3-phosphate synthase gene in cultured carrot cells. *Mol Gen Genet* **232**: 377-382.
- Smith, M.A., Weaver, V.B., Young, D.M., and Ornston, L.N. (2003)** Genes for chlorogenate and hydroxycinnamate catabolism (*hca*) are linked to functionally related genes in the *dca-pca-qui-pob-hca* chromosomal cluster of *Acinetobacter* sp. strain ADP1. *Appl Environ Microbiol* **69**:524-32.
- Stark, G.R., and Wahl, G.M. (1984)** Gene amplification. *Annu Rev Biochem* **53**: 447-491.
- Stark, G.R. (1993)** Regulation and mechanisms of mammalian gene amplification. *Adv Cancer Res* **61**: 87-113.
- Tan, H.M. (1999)** Bacterial catabolic transposons. *Appl Microbiol Biotechnol* **51**: 1-12.
- Thao, M.L., Baumann, L., Baumann, P., and Moran, N.A. (1998)** Endosymbionts (*Buchnera*) from the aphids *Schizaphis graminum* and *Diuraphis noxia* have different copy numbers of the plasmid containing the leucine biosynthetic genes. *Curr Microbiol* **36**: 238-240.

- Tlsty, T.D., Albertini, A.M., and Miller, J.H. (1984)** Gene amplification in the *lac* region of *E. coli*. *Cell* **37**: 217-224.
- van der Meer, J.R., de Vos, W.M., Harayama, S., and Zehnder, A.J. (1992)** Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol Rev* **56**: 677-694.
- van der Meer, J.R., Werlen, C., Nishino, S.F., and Spain, J.C. (1998)** Evolution of a pathway for chlorobenzene metabolism leads to natural attenuation in contaminated groundwater. *Appl Environ Microbiol* **64**: 4185-4193.
- van der Meer, J.R., and Sentchilo, V. (2003)** Genomic islands and the evolution of catabolic pathways in bacteria. *Curr Opin Biotechnol* **14**: 248-254.
- Vasseghi, H., and Claverys, J.P. (1983)** Amplification of a chimeric plasmid carrying an erythromycin-resistance determinant introduced into the genome of *Streptococcus pneumoniae*. *Gene* **21**: 285-292.
- Wagner, A. (2001)** Birth and death of duplicated genes in completely sequenced eukaryotes. *Trends Genet* **17**:237-239.
- Wagner, P.L., Acheson, D.W., and Waldor, M.K. (1999)** Isogenic lysogens of diverse shiga toxin 2-encoding bacteriophages produce markedly different amounts of shiga toxin. *Infect Immun* **67**: 6710-6714.
- Waldor, M.K., and Mekalanos, J.J. (1996)** Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910-1914.
- Wheelis, M.L., and Stanier, R.Y. (1970)** The genetic control of dissimilatory pathways in *Pseudomonas putida*. *Genetics* **66**: 245-266.
- Whoriskey, S.K., Nghiem, V.H., Leong, P.M., Masson, J.M., and Miller, J.H. (1987)** Genetic rearrangements and gene amplification in *Escherichia coli*: DNA sequences at the junctures of amplified gene fusions. *Genes Dev* **1**: 227-237.
- Wolfe, K.H., and Shields, D.C. (1997)** Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**: 708-713.
- Wu, C.H., and Black, L.W. (1995)** Mutational analysis of the sequence-specific recombination box for amplification of gene 17 of bacteriophage T4. *J Mol Biol* **247**: 604-617.
- Wu, C.H., Lin, H., and Black, L.W. (1995)** Bacteriophage T4 gene 17 amplification mutants: evidence for initiation by the T4 terminase subunit gp16. *J Mol Biol* **247**: 523-528.
- Wu, D.G., and Black, L.W. (1987)** Gene amplification mechanism for the hyperproduction of T4 bacteriophage gene 17 and 18 proteins. *J Mol Biol* **195**: 769-783.

- Wu, D.G., Wu, C.H., and Black, L.W. (1991)** Reiterated gene amplifications at specific short homology sequences in phage T4 produce Hp17 mutants. *J Mol Biol* **218**: 705-721.
- Yamaguchi, H., Yamashita, T., Shimizu, H., and Ikeda, H. (1995)** A hotspot of spontaneous and UV-induced illegitimate recombination during formation of lambda bio transducing phage. *Mol Gen Genet* **248**: 637-643.
- Yanisch-Perron, C., Vieira, J., and Messing, J. (1985)** Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.
- Young, D.M., Parke, D., D'Argenio, D.A., Smith, M.A., and Ornston, L.N. (2001)** Evolution of a catabolic pathway. *ASM News* **67**: 362-369.
- Young, M. (1983)** The mechanism of insertion of a segment of heterologous DNA into the chromosome of *Bacillus subtilis*. *J Gen Microbiol* **129 (Pt 5)**: 1497-1512.
- Zabriskie, J.B. (1964)** The role of temperate bacteriophage in the production of erythrogenic toxin by group A Streptococci. *J Exp Med* **119**: 761-780.