Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994. 48:193–222
Copyright © 1994 by Annual Reviews Inc. All rights reserved

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

COMPARATIVE MOLECULAR BIOLOGY OF LAMBDOID PHAGES

A. Campbell
Department of Biological Sciences, Stanford University, Stanford, California 94305

KEY WORDS: antitermination, integrase, recombination in nature, lysogeny, defective prophages

CONTENTS

INTRODUCTION 194
GENETIC FUNCTIONS 196
Repression and Its Control 196
Productive-Cycle Transcription 198
Replication 201
Structure and Assembly of Virions 201
Lysis Genes 203
Genes for Insertion and Excision 204
Accessory Genes 207
DEFECTIVE LAMBDOID PROPHAGES 207
RECOMBINATION AND PHAGE PHYLOGENY 209
Comparative Genome Structure 209
Genesis of the New Combinations 212
Molecular Phylogeny 215
Selection for Polymorphism 215
Accessory Genes 216
SPECIATION IN PHAGES 217

ABSTRACT

Lambdoid phages are natural relatives of phage λ. As a group, they are highly polymorphic in DNA sequence and biological specificity. Specificity differences have played a key role in identifying the specific sequences recognized by the N and Q antitermination proteins, the initiator O for DNA synthesis, the terminase system Nu1-*A* for cutting DNA during packaging, and the *cI* repressor protein. Variations that go beyond specificity differences are seen in packaging mechanism (headful in P22, specific cutting in lambdoid

0066-4227/94/1001-0193$05.00

Annual Reviews
www.annualreviews.org/aronline

coliphages), in early control (terminator protein and phage-independent antitermination in HK022, phage-specific antitermination in λ), in repression control (antirepressor operon in P22, absent in other lambdoid phages) and murein-degrading enzymes (transglycosylase in λ, lysozyme in other lambdoid phages). Sequence comparisons indicate that recombination among lambdoid phages is frequent in nature.

INTRODUCTION

Starting with the early studies on the nature of repression (57, 59), the temperate coliphage λ has served as a model for the fields of gene regulation and temporal programming of gene expression. The λ phage has been useful because, following infection, λ development can proceed along two alternative pathways. Some cells enter a productive cycle, in which phage DNA replicates autonomously and is packaged into progeny phage particles, which are then liberated by lysis. Other cells survive infection to become lysogenic and harbor the phage DNA inserted into the chromosome as a prophage, which remains transcriptionally quiescent for genes of the productive cycle.

Natural relatives of λ have been isolated from various sources (11, 31, 53, 63, 64). Most of them grow on *Escherichia coli*, but a few (such as P22, L, and LP-7) come from *Salmonella typhimurium*. These phages have the same gene order as λ and can form viable recombinants with it. Most of them differ from λ in the specificity of genetic determinants for repression, integration, regulation, packaging, and/or cell surface recognition. Such specificity differences have played a key role in studies of λ development. For example, a hybrid phage whose entire genome came from λ, except for a few kilobase pairs from phage 434, forms lysogens that are sensitive to λ infection and immune to 434 infection and can undergo a productive cycle on infecting a λ lysogen but not a 434 lysogen; this observation implies that not only the determinant for repressor (*cI* gene) but also the determinants of repressibility (operator sites), and any additional *trans*-acting factors specific for the same sites (such as the *cro* gene), lie within that DNA segment (60). Later helix swap experiments pinpointed repressor specificity to a small segment of the protein (101). And the fact that the repressed promoter of a prophage is inactive, even when the cell is productively infected by a heterospecific phage, demonstrated that the positive control factor gp*N* causes antitermination of transcripts coming from the major early promoters rather than downstream initiation of new transcripts (69).

Within each block of genes depicted in Figure 1, the variation among lambdoid phages indicates which features of sequence or structure have been highly conserved. Beyond providing information deducible from systematic gene-bashing of λ itself, the results sometimes reveal alternative solutions to

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

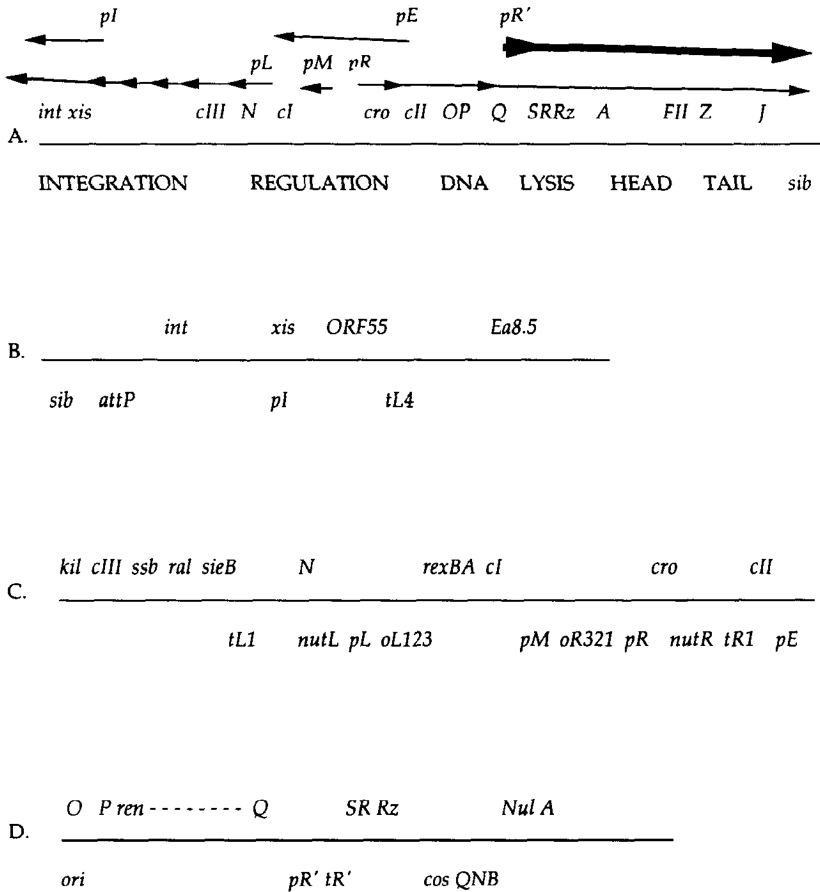

Figure 1 (*A*) Map of λ prophage, showing major gene clusters, some of the genes they contain, and major transcripts. Antiterminated transcripts contain arrowheads at terminator sites. Gene *Q*, the late regulator, appears between clusters. DNA-recognition sites are not shown, except for *sib*. (*B*) Integration genes (*above line*) and recognition sites (*below line*) discussed in the text, plus upstream DNA. (*C*) Early regulation genes and recognition sites discussed in the text. (*D*) Replication, lysis, and packaging genes discussed in the text. The light dashed lines between *ren* and *Q* are accessory genes of unknown function, named for the molecular weights of their predicted protein products: 146, 290, 57, 60, 56, 204, 68, and 221.

common problems. The light shed on function by surveys of natural diversity is a recurrent theme in this review. Here, I discuss some recent progress in λ molecular biology but emphasize only those aspects relevant to comparative studies.

Annual Reviews
www.annualreviews.org/aronline

Beyond enriching our understanding of λ, the study of lambdoid phages has raised questions of its own. How has the variation among different types arisen and how has it been maintained? How are these phages related to one another, and are they better understood as individual lines subject to variation or as members of a common gene pool subject to frequent genetic exchange? These questions are also addressed herein. Three previous reviews (18, 20, 25) have covered some of the evolutionary issues.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

GENETIC FUNCTIONS

Repression and Its Control

In bacteria lysogenic for λ, productive cycle genes are turned off because the repressor binds to two operators (*o*L and *o*R). Each operator consists of three oligonucleotide sites separated by short spacers, and each site has approximate two-fold symmetry. The rightward operator *o*R controls both rightward transcription of productive-cycle genes from promoter *p*R and leftward transcription of *cI* from promoter *p*M (Figure 1). Rightward transcription is repressed by the cooperative binding of repressor to sites *o*R2 and *o*R3, which also stimulates *cI* transcription. The first gene product of the rightward transcript (Cro) binds the same sites as the repressor. Cro bound to *o*R3 (for which it has the highest affinity) prevents *cI* transcription. At higher Cro concentrations, *o*R2 and *o*R1 are filled, and rightward transcription shuts off as well. A lucid, well documented summary of λ regulation is available (78).

In each infected cell, the concentration of the product of gene *cII* determines whether the cell enters the productive or the lysogenic cycle. This gene product stimulates leftward transcription of *cI* from promoter *p*E within *cII*. The CII protein is highly unstable in vivo, and CIII inhibits its proteolysis. Those cells that achieve a high concentration of CII make repressor and enter the lysogenic state. Then, once repression is established, the *p*R promoter is shut off and CII disappears from the cell. A sufficient CII concentration can only be attained soon after infection, before Cro accumulates and reduces CII production by repressing *p*R. The effect of Cro at *p*M has at most an ancillary role in this process; its main importance is probably in reinforcing the directive to enter the productive cycle when a lysogen is derepressed.

All lambdoid phages have an arrangement of regulatory elements similar to that in λ. At least 10 distinct repressor specificities have been reported. Six *cI* genes that have been sequenced are all related, indicating common ancestry (17, 25).

MUTATION TO VIRULENCE The selective forces that preserve the ability to achieve a stable lysogenic state (and therefore explain why it pays a phage to

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

be temperate) are incompletely understood. Whatever the advantages of temperateness, they must be long range. A single particle of a virulent mutant (one able to infect and lyse a lysogen, sometimes because the operator sites have mutated to forms unable to interact with repressor) can multiply in and cause destruction of a whole culture of lysogenic cells. Thus, the virulent mutant clearly has the short-term advantage, and lysogeny can be stable only under conditions where virulent mutants rarely infect, such as low cell density, low rate of mutation to virulence, or some other safeguard against infection [for example, that seen in phage ε15 of *Salmonella typhimurium,* in which the cell surface is so altered by prophage-coded enzymes that ε15 can no longer attach (98)]. In λ, at least three mutations, two in the right operator and one in the left, are needed for virulence (75). Lambdoid phage HK022 (whose left operator controls no functions essential for productive infection) requires two mutations, one in *o*R1 and one in *o*R2 (23). The single mutants are still temperate because of an additional repressor binding site *o*FR (operator far right) located downstream from *cro,* 267 bp from *o*R3. *o*FR mutants have no obvious phenotype, but *o*R1 *o*FR and *o*R2 *o*FR double mutants are virulent. The mechanism of *o*FR action is unknown.

Salmonella phage P22 has an apparently different way of avoiding single-step virulence. The *sieA*$^+$ gene (which causes superinfection exclusion) prevents single-step mutations in the *mnt*-controlled operator from causing virulence (23, 94). P22 differs from the lambdoid coliphages in having a second operator-repressor system (Ant-Mnt) that regulates the primary repressor. The *ant* and *mnt* genes are contiguous but distant from the gene for the primary repressor and its operator sites. Ant protein complexes with and neutralizes the repressors of lambdoid phages, including P22. Mnt represses Ant synthesis in P22 prophage.

REPRESSOR BINDING Although the repressors from different lambdoid prophages have diverged in sequence and specificity, the arrangement of binding sites and mechanism of binding appear to be well conserved. The λ repressor binds to the operator sites as a tetramer or octamer (87). X-ray diffraction studies of repressor-operator complexes show that both λ and 434 repressors make similar contacts near the outer edges of the pseudosymmetric operator sites and less symmetrical contacts near the center, where the specificity determinants lie (1, 9). Like λ repressor, HK022 repressor binds cooperativity to *o*R1 and *o*R2, with a cooperativity coefficient of 1600, compared to 50–300 for λ (22). In stabilizing the lysis/lysogeny decision, the stronger cooperativity in HK022 may counterbalance the weaker binding of HK022 Cro compared to λ Cro.

REPRESSOR SYNTHESIS Activation of *p*M transcription by repressor bound at *o*R2 appears to require contact between RNA polymerase and a cluster of acid

Annual Reviews
www.annualreviews.org/aronline

groups in the repressor. This is true not only for λ but also for five other lambdoid phage repressors whose tertiary structures have been predicted from amino acid sequence (17). An unusual feature of *p*M is that the initiating AUG in the *cI* gene lies at the 5′ end of the transcript, with no upstream ribosome-binding site. *cI* also has a downstream sequence that improves translatability, which in turn is inhibited by ribosomal protein S2. The absence of a leader and presence of the downstream sequence allow more efficient complexing of the *p*M message with S2-depleted 30S subunits and initiator tRNA (88). Toe-printing suggested the remarkable possibility of direct complex formation between the transcript and undissociated 70S ribosomes (6). Whatever the functional significance of these usual properties may be, they are not universal. The 434 *p*M transcript has a leader sequence with a good ribosome-binding site (17).

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

cII SPECIFICITY Whereas many different repressor specificities are found among lambdoid phages, the CII proteins of different phages are frequently but not always interchangeable. CII activates three promoters, all functional in lysogenization, widely separated in the phage genome. *p*E controls repressor synthesis, *p*I controls integrase synthesis, and *p*AQ makes an antisense RNA that inhibits lytic development. Coevolution of all three elements to a new specificity should be impeded by natural recombination that occasionally associates *cII* genes with new *p*Is and *p*AQs. Nevertheless, the CIIs of λ and P22 are not fully exchangeable, although both proteins recognize the consensus CII-specific target $TTGCN_6TTGC$ at −35 (56).

Productive-Cycle Transcription

EARLY TRANSCRIPTION Following infection, RNA polymerase initiates leftward and rightward transcription from the strong promoters *p*L and *p*R, respectively. The immediate products of this transcription are *N* message on the left and *cro* message on the right. gp*N* has the remarkable ability to alter RNA polymerase so that it overrides ordinary termination signals like *t*L1 and *t*R1 and allows transcription to proceed throughout almost the entire λ genome. Because *cII* and *cIII* are immediately downstream of *t*R1 and *t*L1, gp*N* promotes lysogeny (as indicated by the fact that limitation of gp*N* supply, as in a partially suppressed amber mutant, shifts the system toward lysis), but its action is essential for the productive cycle as well. The rightward transcript includes genes *O* and *P*, the replication origin (within *O*) and the late gene activator *Q*.

The antiterminating action of gp*N* requires the presence in the transcript of a special sequence, *nut* (*N*-utilization site). At the *nut* site, gp*N* and host transcription factor NusA bind mRNA to form an RNA loop with the RNA polymerase downstream from *nut*. The gp*N*-NusA complex by itself is unstable

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

and only influences termination close to *nut*, but in the presence of other factors (NusB, ribosomal protein S10, and the translation factor NusG) the complex endures long enough to override terminators many kilobases downstream (70).

All *nut* sites of lambdoid phages have two sequence components (separated by 8–12 bp): a nonamer, *boxA*, resembling the consensus CGCUCUUUA (exactly realized in P22) and a potential stem-loop, *boxB* (30). The role of *boxA* in antitermination is not phage specific and in fact may be negative. When *boxA* is deleted, antitermination no longer requires NusB, as expected if *boxA* complexes with some inhibitor and NusB neutralizes that effect (77). The fact that NusB is required for formation of stable antitermination complexes in a purified in vitro system suggests that it may have some other activity as well, or that the inhibitor is one of the components of the purified system. The *boxB* loop, most of whose sequence is required for antitermination, is phage specific (32, 65). Different lambdoid phages have unique or overlapping specificities. gp*N* λ and gp*N*21 work only with their cognate *nut* sites, whereas gp*N* of P22 accepts either one. The *N* genes of these lambdoid phages have no significant sequence homology except at the 3′ end, most of which is not needed for antitermination (39). The positions of the other relevant elements (*p*L, *nut* L, *t*L1, *nutR*, *t*R1) are also conserved, although the sequences have diverged extensively (38).

Of the phages studied so far, lambdoid phage HK022 is the only one that has a gene, *nun*, similarly positioned to λ *N* and with some 3′ homology to it that does not play the same role in antitermination as gp*N*. Nun provokes premature termination of transcription slightly downstream from the *nut* sites of λ. The *nun* gene is expressed even in the prophage state (by readthrough from *p*M) and renders HK022 lysogens resistant to λ infection by truncating the early λ transcripts (74). The Nun and gp*N* λ targets are not fully congruent. Deletion of *boxA* and *boxB* destroys responsiveness to both factors, but deletion of *boxA* alone reduces responsiveness only to Nun (7). Some substitution mutations in the *boxB* loop impair responsiveness to gp*N* λ but not to Nun. Certain mutations either in *nut* sites or in host *nus* genes can reverse the effect of Nun so that it causes antitermination rather than termination (82).

Nun is not required for antitermination of HK022 transcripts, but antitermination does occur downstream of a specific 40-bp sequence in the transcript, but neither phage functions nor Nus proteins are needed. Host mutations that suppress antitermination map in a core subunit of RNA polymerase, suggesting that the polymerase might by itself switch to a configuration where normal terminators are no longer recognized (73).

HK022 studies should further illuminate the mechanism of antitermination, and our understanding of the full circuitry of HK022 and of the reasons(s) for antitermination in λ should develop hand in hand. Given the presence of terminators *t*L1 and *t*R1, antitermination is clearly needed. But why should λ

Annual Reviews
www.annualreviews.org/aronline

preserve terminators only to override them? Some possible explanations include (40):

1. Antitermination allows a quicker response time than simple repression, because a polymerase may remain poised at a pause site until the antiterminator appears. In λ, this is easier to see for late transcription than for early transcription.
2. During evolution, foreign DNA segments were incorporated into λ operons, along with their terminators.
3. Transcription from weak promoters downstream of the *nut* sites can be arrested at terminators that are read through by transcripts from *p*L and *p*R. The existence of such weak promoters is well documented. For example, the conserved and highly efficient terminator *t*L4 stops transcription from a promoter a few hundred nucleotides upstream (5). As this promoter is expressed in the prophage, readthrough into *int* and *xis* might destabilize a lysogen. Whether this promoter plays any role in λ biology is unknown.

Two additional reasons are possible:

4. The use of antitermination stabilizes the lysogenic condition. Although a chance escape from repression can produce *p*L or *p*R transcripts, in the absence of gp*N* these will stop at *t*L1 or *t*R1 rather than express genes that cause excision, replication, or cell death.
5. Antitermination permits control of *int* RNA stability. When an infected cell decides to become lysogenic, *int* is transcribed from the CII-activated promoter *p*I. In a cell that is uncommitted or destined to lyse, it is transcribed from the major leftward promoter *p*L. Whereas *p*I transcripts are fairly stable, *p*L transcripts are degraded, because of RNase III cleavage of a downstream stem loop (*sib* site) followed by 3′ → 5′ exonucleolytic degradation. *p*I transcripts terminate before the complete RNase III site is transcribed. The differential instability of *p*L transcripts thus requires readthrough of a normal terminator. No comparable effect is known in the *p*R transcript.

These reasons are not mutually exclusive. The conservation of terminator function (though not precise sequence) of *t*L4 strongly supports reason 3, but reason 3 may be a secondary consequence of reason 2. If stabilization of lysogens (reason 4) is important, it is not obvious how HK022 achieves stability. In principle, a lysogen could be stabilized by a termination factor expressed in lysogens (like Nun) rather than an antitermination factor made after repression is lifted (like gp*N*λ). However, no evidence suggests that Nun causes termination in HK022 or loses activity after induction.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

LATE TRANSCRIPTION Among the early gene products made from the *p*R transcript is the late gene activator gp*Q*. As infection proceeds, the accumulation of Cro slows early transcription, and gp*Q* stimulates transcription of genes for lysis and virion formation. This stimulation results from antitermination of transcription from the constitutive promoter *p*R′. Some *p*R transcripts continue through *p*R′, but the resulting rate of *Q*-independent transcription of late genes is much less than the *Q*-dependent transcription from *p*R′.

In absence of gp*Q*, transcription from *p*R1 proceeds for 16 nucleotides, then pauses for several minutes before proceeding to the terminator at +194. This pause site, plus the downstream sequence, define a Q-utilization site (*qut*). Unlike *nut* sites, *qut* sites only function when close to the promoter; at least in phage 82, transcription from an upstream promoter is not antiterminated (43).

Pausing is necessary but not sufficient for antitermination. The three phages λ, 82, and 21 all have pause sites (21 has two, at +15 and +25), but each gp*Q* is specific for its own *qut*. The three *qut* sites have little homology with one another (48, 103).

Replication

Autonomous replication of λ requires two gene products of the *p*R transcript (gp*O* and gp*P*) initiating bidirectional replication from an origin within gene *O*. Replication also requires activation *in cis* by rightward transcription. This transcription normally comes from *p*R and passes through *ori*, but artificial introduction of a *lac* promoter at various sites shows that the critical transcription is downstream from *ori* (50). The bulk of gp*O* is metabolically unstable, but about 20% is apparently stabilized by its location in replication complexes (99). As replication proceeds, gp*O* remains bound to the origin, but also to the growing points of the DNA (85). After a period of theta-form replication, most molecules switch to a rolling-circle mode, where multigenomic lengths of double-stranded DNA are generated.

At initiation, gp*O* specifically recognizes both *ori* and gp*P*, which binds to the host *dnaB* helicase (105), and chain elongation requires the full complement of host-elongation proteins. The common lambdoid phages have several different replication specificities. gp*O* is specific both for its cognate *ori* (in its N terminus) and gp*P* (in its C-terminus) (41). The *O* genes of different lambdoid phages are recognizably homologous. In phage P22, gene *P* is replaced by a *dnaB* homologue, gene 12 (4). Host DnaB is not required for P22 replication, but gene 12 is essential, even in a $dnaB^+$ host (13).

Structure and Assembly of Virions

HEAD STRUCTURE After replication is underway, synthesis of virion components is initiated, and these are assembled by parallel pathways into proheads

Annual Reviews
www.annualreviews.org/aronline

and tails. DNA packaging into proheads proceeds from multigenomic tails of rolling circles, or from multimeric circles produced by recombination. Packaging initiates at a *cos* site, where DNA is cut by the terminase complex (phage proteins gp*A* and Nul). DNA to the right of the cut is then pumped into the prohead, and finally the next *cos* site is cut to produce the right end of the packaged DNA. Packaging is processive along three to four genome segments. Once the prohead shell (whose major component is gp*E*) is filled, it expands and assumes a more obviously icosahedral shape, and a second major protein, gp*D*, is added. Tails are then attached to the completed heads.

All the head genes lie at the left end of the λ map, followed by the tail genes. This functional clustering extends to a finer level. Genes whose products interact with one another, such as *Nul* and *A*, are generally adjacent (and in this case also adjacent to their target site, *cos*), and the last gene of the head cluster (*F*II) encodes a protein at the head-tail junction, which interacts with the product of the first tail gene, *Z*.

Different phages have sometimes diverged so far that heterologous products no longer interact productively. For example, phages 21 and λ both have 10 head genes arranged in the same order, with about 50% amino acid identity in each, but 21 fails to complement λ mutants in head genes other than *D* (89). Even phages whose head genes are not appreciably homologous with λ (such as P22 and HK97) share the same general organization (25).

The specificity of λ and 21 terminases has been used to explore the mechanism of DNA cutting. The *cos* site can be dissected into at least three components; an initiation segment (*cosB*) on the right, a nicking site (*cosN*), and a termination site (*cosQ*) on the left (29). (After cutting, *cosQ* lies near the right end of one molecule and *cosB* near the left end of another.) The *cosB* segment includes an IHF binding site and three binding sites for the Nul subunit of terminase. Within one phage, these sites are similar to one another; but the λ sites differ from those of phage 21. Some λ *Nul* mutations can overcome the packaging defect caused by mutations in the binding sites (28, 90). The results of packaging prophage DNA from tandem triple lysogens show that specific binding at *cosB* is required only to initiate packaging in the first of a series of processively packaged copies. Once cutting is initiated at a *cosB* recognizable by the available gp*Nul*, it proceeds in a rightward direction through adjacent copies, regardless of their *cosB* specificity (36). Mutations in *cosQ* allow DNA entry into the prohead but prevent its final cutting (29).

The 21 analogue of Nul inhibits λ terminase both in vivo and in vitro. λ mutants resistant to inhibition overexpress *Nul* (through alteration of its ribosome binding site), indicating that inhibition is competitive (58).

Thus, in the coliphages λ and 21, the basic packaging mechanism is similar, the specificity difference lying in the binding of Nul to *cosB*. The lambdoid *Salmonella* phages superficially follow an entirely different plan. DNA is not

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

cut at specific *cos* sites but rather is processed into headful lengths. However, the two phage types share some commonality, which perhaps indicates common ancestry of their head genes. In both cases, packaging initiates at a site (λ *cos* or P22 *pac*) located near the terminus of a cluster of virion genes, and continues processively and undirectionally through several copies (36, 94). In P22, headful packaging creates a terminal redundancy about 8.7% of the genome length; in L, the fraction is 7.2% (91).

Phage HK97 has the remarkable feature (thus far unique among lambdoid phages) of extensive covalent linkage of major head protein subunits within the virion (25).

TAIL STRUCTURE The λ tail consists of a long hollow tube (through which DNA is injected during infection) terminating in a single fiber encoded by gene *J*. In λ as isolated from nature (Ur-λ), the sides of the tail are adorned with bent tail fibers encoded by two genes, *stf* (side tail fibers) and *tfa* (tail fiber assembly) (52). These genes lie close to and downstream from *J*. The common laboratory strains of λ have a frameshift in *stf* that leads to a truncated, functionless peptide product. In laboratory strains, contact between gp*J* and the host protein LamB is essential for phage attachment and therefore for plaque formation. In Ur-λ, *stf* fibers can cause stable attachment to the cell surface (probably to the major outer membrane protein OmpC).

Most lambdoid coliphages have virion structures similar to λ's structure. At least three different host ranges attributable to different attachment specificities are known (53). Phage PA-2 uses OmpC as its primary receptor (83). The lambdoid *Salmonella* phages have a radically different morphology: short stubby tails with a base-plate, from which projects a single fiber of the gp*J* analogue used for attachment (94).

Lysis Genes

Three adjacent λ genes *S*, *R*, and *Rz* are dedicated to lysis (104). These are the first three genes of the late operon transcribed from *p*R′. The S protein (a holin) forms holes in the cytoplasmic membrane, thereby collapsing the membrane potential, killing the cell and allowing the endolysin (*R* protein) to reach and cleave its substrate, the rigid murein layer. Cleavage occurs between N-acetylglucosamines (the same bonds cleaved by true lysozymes), but the gp*R* cleaves by transglycosylation rather than hydrolysis. Other lambdoid phages (21, PA-2, P22) have true lysozymes. The mechanism of *Rz* action is unknown. It is needed for lysis under special conditions, such as high Mg^{2+} concentration. λ lysates contain an activity for hydrolysis of the short D-amino acid cross-bridges between N-acetylglucosamine chains (95). This function has been suggested for *Rz*, on no direct evidence.

The *S* gene produces two protein products from different translational start

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

sites. The shorter peptide has holin activity; the larger one inhibits it. This balance between action and inhibition does not seem to play a significant role in the timing of lysis, but may stabilize lysogens against the potentially lethal effect of chance readthrough of the first terminator in the late operon (12). Although the *S* genes of λ and 21 have no detectable homology, the dual start motif is conserved in both phages.

Genes for Insertion and Excision

In those infected cells destined to survive as lysogens, circularized λ DNA is inserted into a specific chromosomal site by the phage-coded integrase (100). The minimal substrates for the reaction are a 21-bp segment of host DNA and a supercoiled phage DNA including about 240 bp of specific sequence flanking the crossover point. In the reaction, the N termini of integrase molecules bind strongly to arm sites (which lie close to the extremes of the 240-bp *attP* DNA), and a host protein (integration host factor) binds between the arm sites and the crossover point and bends DNA so that the catalytic sites on the C termini are positioned close to the crossover point (62). In the 21 bp surrounding the crossover point, Int recognizes weakly binding core sites symmetrically flanking the crossover point, and the crossover proceeds through formation of a Holliday crossbridge between strands of like polarity, followed by rightward branch migration through a 7-bp overlap segment and resolution by a crossbridge between the other two strands. The intermediates in the λ pathway are transient and have been trapped only by use of artificial substrates; with phage HK022, the resolution step is apparently slower, and DNA molecules with a single crossbridge are detectable in the reaction mix (71).

When repression is lifted in a lysogenic bacterium, phage DNA is excised from the chromosome. The excision reaction requires an additional phage-specified protein (Xis), which binds to λ DNA between the left arm and the core. Of the two Xis-binding sites, one can be occupied by the bacterial protein FIS in place of Xis (96). Excision is superficially the reverse of insertion, but it is not a true chemical reversal. In a reversal of the insertion reaction, branch migration would proceed leftward rather than rightward, but in both insertion and excision, branch migration is rightward (72).

RNA polymerase transcribes the *int* gene from two alternative promoters (Figure 1). In a cell that commits to lysogeny, gp *cII* activates a promoter, *p*I, whose transcript expresses only *int*, not the upstream gene *xis*. In an infected cell or a derepressed lysogen, *int* and *xis* are transcribed coordinately from *p*L. In a derepressed lysogen, *int* and *xis* are both expressed; however, in an infected cell, *int* expression from the *p*L message is strongly reduced by 3′ exonucleolytic degradation of transcripts cleaved at the *sib* site by RNase III. Because *sib* lies downstream from the crossover point, the insertion event separates *sib* from *int;* hence, the prophage transcript in a derepressed lysogen is stable (34).

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

SITE RECOGNITION The *int* genes of all lambdoid phages are related to one another, indicating common ancestry, but site-specific recombinases related to *int* are also distributed among nonlambdoid phages, bacterial plasmids, and eukaryotes (3, 27). Xis protein is small, and heterospecific *xis* genes have little homology, although some peptide motifs are conserved (84). Some lambdoid phages, such as λ and 434, insert at the same chromosomal site, and all components of the system are functionally interchangeable. In contrast, λ and 21 insert at different sites; neither Int nor Xis is interchangeable; and there is no detectable homology in arm recognition sites, core recognition sites, or overlap segments (66, 84). At least nine different types of lambdoid phages can be distinguished based on insertion sites.

Although the core recognition sites and overlap segments differ, the lambdoid phages appear to conserve the same spacing and probably undergo the same molecular gymnastics during crossbridge formation and branch migration. The spacing between arm sites and core sites is rigidly conserved among conspecific phages such as λ, 434, and HK022 (which inserts at a different site but has the same arm-site recognition), as might be expected for proper alignment of core sites with the C-termini of integrase molecules (102). However, for some nonlambdoid phages (such as HP1), the minimal *attP* site is more than 400 bp (compared to 240 for λ), suggesting that the spacing is different (51). Indirect evidence suggests a different spacing for 21 as well (84). Thus more than one pattern of DNA looping has probably developed from a common ancestor.

Phage pairs that have diverged far enough to change some, but not all, aspects of site specificity are especially interesting. λ and HK022 insert at different sites and have distinct but overlapping specificities of core site recognition (71, 102). Their *xis* genes and the 5′ ends of their integrase genes are very similar; accordingly, their Xis proteins are interchangeable and their arm site recognition is identical. Phage 21 and defective element e14 insert at the same site, but appear to recognize different sequences both at the core and in the arms (21, 84).

Does core-site recognition correlate with integrase phylogeny? Those integrases that most closely match λ integrases lie along two branches of the tree (84). λ/434 integrases (equivalent or almost identical) are most closely related to HK022. On the other branch lie phages that insert into structural genes for host proteins (21 and e14 in *icd,* and Atlas in an open reading frame of unknown function). The complete sequence of 21 *int* is available, and partial sequences for e14 and Atlas. All these elements include in their *attP* DNA an imprecise duplication of the host-gene sequence 3′ from the crossover point, so that the lysogen produces a protein product from the chimeric gene that extends across *attL;* the larger size of 21 *attP* compared to λ may reflect the positioning of the arm-recognition sites distal to the duplication. Atlas is present (either as

Annual Reviews
www.annualreviews.org/aronline

Table 1 Core *att* sequences of some lambdoid phages

Insertion site	Phage	Sequence[a]	Reference
Intergenic	κ*attB*	CtgcTTttTtatActAActtg	Sec 19
	attP	CAgCTTttTtatActAAGtTG	
	HK022*attB*	gcaCTTTaggtgaaAAAGgtt	See 19
	attP	atcCTTtaggtgaatAAGttg	
Intragenic tRNA	p22*attB*	CagCGCATTcgtAATGCGaag	See 19
	attP	CTTataATTcgtAATgcgAAG	
	DLP12*attL*	CaACGACcTtctAaGTCGTgG	67
	attR	CtAgGAacTtctAaGTCgTgG	46
ORF	21 *attB*	gAgATGatGCtGCgcCATaTg	See 19
	attP	tATATcctGCtGCgccATATg	
	Atlas *attB*	aAcaTgtCAgtgTGgtAcaTg	See 19
	attP	aAacTACCAgtgTGGTAcaTg	

[a] Sequences are aligned around centers of symmetry. For each sequence, those bases are capitalized that are symmetrical (complementary) around the center. Symmetrical consensus sequences have been proposed (19). The DNA that is identical between *attB* and *attP* is underlined in the top sequence.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

an active or a defective prophage) in many natural *E. coli* strains (not in K-12) (93). A third pair (minimally related to λ or 21 but closely related to each other) comprises P22 and the defective coliphage DLP12 (67). These phages insert into the anticodon loops of tRNA genes and precisely duplicate the 3′ ends of these genes.

Symmetrical consensus core-recognition sequences have been proposed for all these phages except DLP12 (19). For λ, HK022, and 21, evidence clearly demonstrates that some of the consensus positions are important. Whereas the consensus sequences for the various phages differ, the members of each pair (HK022 vs λ, 21 vs Atlas, DLP12 vs P22) are somewhat similar, but not enough to suggest specific routes whereby specificity changes have taken place (Table 1).

REGULATION In λ, the gp*cII*-activated promoter p*I* overlaps the 5′ end of the *xis* gene, so that the only full open reading frame within the transcript is the *int* gene. Promoters with the CII-recognition sequence are similarly located in 434 and HK022 (5, 102). Neither phage P22 nor phage 21 has such a promoter. It is not known whether or how *int* transcription is regulated in these phages. Downstream sequences that could function as *sib* sites are present in 434, HK022, 21, P22, and DLP12 (5, 84, 102).

Annual Reviews
www.annualreviews.org/aronline

Accessory Genes

Some genes in lambdoid phages are not required (and frequently not even noticeably beneficial) for either the productive or the lysogenic cycle (26). Some of these confer properties of plausible natural advantage, such as ability to exclude possible competitors (*rexAB* and *sieB* in λ, *sieAB* in P22, *nun* in HK022) or increased survivability of lysogens in the presence of serum [*bor* in λ (8)]. Typical accessory genes are present in some, but not all, lambdoid phages. A cluster of accessory genes (starting with *ren*, which provides resistance to Rex exclusion) lie in the interval between *P* and *Q*. The DNA between *J* and *sib* also plays no known role in lysis or lysogeny and includes several genes such as *lom*, *stf*, and *tfa* and other unidentified open reading frames.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

DEFECTIVE LAMBDOID PROPHAGES

Some natural bacteria are lysogenic for one or more lambdoid prophages. Prophage genes such as *rex* or *lom* may confer a selective advantage on the host, but in general most prophage DNA constitutes junk whose eventual elimination is expected through random drift, slightly aided by selection against the small expense of replicating the extra DNA and occasionally losing a cell through spontaneous lysis. Elimination by clean deletion would leave no trace, but frequently prophages undergo partial deletion, leaving telltale remains. Thus, Southern hybridization of natural enterobacteria against a λ probe shows extensive dispersion and length polymorphism of λ-related sequences (2).

The K-12 strain of *E. coli* harbors, in addition to λ itself, four defective lambdoid prophages: DLP12, Rac, e14, and Kim. DLP12 (previously called *qsr'*) is about 25 kb long, compared to 49 kb for λ (67). The first two kb have been sequenced. An integrase gene overlaps the tRNA gene *argU*, into which DLP12 is apparently inserted. Upstream (rightward) of *int* is a portion of the *xis* gene—both *int* and *xis* are about 70% identical in amino acid sequence to *int* and *xis* of P22. Next to these are DNA segments closely related to the *redX* and *P/ren* genes of λ, followed immediately by an IS*3* insertion. Farther to the right come analogues (*qsr'*) of the λ *Q*, *S*, and *R* genes, followed by an IS5 insertion, *nmpC*, and *cos* (67, 80). Still farther on is a 151-bp sequence (within gene *appY*) 84% identical to part of λ ORF 194 (which lies near the right end of the λ prophage) and DNA upstream of *ompT* that includes a *sib* site similar to that of phages 21 and P22 and 38 bp identical to part of *argU* that could include the core region of an *attR* site. Thus, DLP12 contains phage sequences in their proper order, some related to known lambdoid phages, two IS insertions, and several major deletions.

Annual Reviews
www.annualreviews.org/aronline

DLP12 surfaced as the source of pseudorevertants of λ *Q* mutants and later of *cos* mutants, generated by recombinational replacement of mutant genes of an infecting λ by their analogues in DLP12 (37, 92). Its homologies with λ prophage allow its use as one of the two crossover points generating replacements of some of the prophage DNA with segments of the bacterial chromosome (80). In both *cos* rescue and replacements within the prophage, at least one of the two crossover points lies in homologous DNA, but sometimes the second recombination was apparently illegitimate.

Southern hybridization showed that the *argU*/DLP12 junction is present not only in K-12 but also in some distantly related bacteria such as *Salmonella boydii* (67). It is not clear whether this junction has been preserved over evolutionary time (which would suggest a useful function to the bacterium) or has been regenerated by a recent lysogenization.

The Rac prophage was discovered by its ability to promote recombination in UV-stimulated cells (68). This phenotype results from the presence of a repressor gene (SOS-activable, like that of λ) controlling a recombinase, *recE* [similarly located but different in sequence from λ's *red* genes (42, 61)]. The prophage (about 27 kb) is flanked by intact *att* sites and can apparently be excised by its integrase system. A segment including *att* and *recE* can recombine into λ, replacing its λ analogues and generating a phage called λ *reverse* (44). A replication origin is present, at the same location as in other lambdoid phages.

The 14-kb element e14 was detected through its occasional excision from the chromosomes of SOS-induced cells (45). Subsequent study showed that SOS stimulates the production by e14 of a factor (presumably an integrase) that can catalyze site-specific recombination between e14 sites cloned into a plasmid (15). Part of a candidate integrase gene has been sequenced. It does not lie in the usual position (next to the insertion site) and would require three frameshifts to express integrase. The DNA adjacent to the insertion site has no integrase homology (84), and no other homology to λ genes has been found in e14. It contains several genes discovered by their interactions with other elements, a gene (*lit*) affecting late transcription of phage T4, an invertible segment (*pin*) similar to the *hin* segment controlling phase variation in *S. typhimurium,* and a restriction gene, *mcrA* (15).

The Kim prophage carries a set of *QSR* analogues (different from those in DLP12) flanked by λ homology that extends beyond *cos* (35). Upstream of the *Q* analogue are a group of genes (*dic*) that interact with the cell-division system and are thought to be derived from the *cro* (*dicC*), *cI* (*dicA*), and *kil* (*dicBDEF*) genes of some lambdoid phages (14). The *dicC* and *dicA* genes are related in sequence to the P22 genes encoding Cro and repressor, respectively (10). The *dicF* gene, which specifies a small RNA product, is closely related to a sequence at the corresponding position of the Rac prophage (CC Chu, L Satin, AJ Clark, M Faubladier & J-P Bouché, personal communication).

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

RECOMBINATION AND PHAGE PHYLOGENY

Comparative Genome Structure

Generally, many parts of the DNA sequences of a typical pair of lambdoid phages such as λ and 21 are not recognizably homologous, yet the arrangements of both the major gene clusters and sites and signals within a cluster are conserved. The relationship between *E. coli* and *S. typhimurium* is superficially similar; gene order is conserved in the face of approximately 15% sequence divergence (81). For bacteria, the favored rationale for conservation of order is selective. For example, any rearrangement that disrupted the preferential orientation of transcription units in the direction of replication-fork movement might be a disadvantage (16). Lambdoid prophage insertion sites are located and oriented nonrandomly on the *E. coli* chromosome, for reasons that could be selective (21).

With lambdoid phages, selection for preservation of regulatory circuitry might be adequate to conserve genome organization. However, another factor intervenes. Very probably, lambdoid phages have repeatedly encountered and recombined with one another in Nature, thus placing all such phages in a common gene pool (18). Interstrain recombination (frequently causing substitutions of DNA segments less than 1 kb in length) is conspicuous among natural lines of *E. coli* but not between *E. coli* and *S. typhimurium* (33). The ability to generate viable recombinants might thus select for maintenance of gene order in the phage (but not the bacteria).

Awareness of the recombinational origin of natural lambdoid phages began with analysis of molecular heteroduplexes between complementary strands of different phages, culminating in a comprehensive set of pairwise comparisons among six lambdoid phages (55). Some general features are evident:

1. Phage pairs form homoduplexes in some segments, not in others. For example λ and 434 match in nine different segments of the linear chromosome, separated by eight segments of heterology.
2. Few segments are universally conserved. Among the six phages λ, 434, 82, PA-2, 21, and 424, the only places where homoduplex formation was observed throughout were the tail cluster and the *Rz* gene. Elsewhere, each pair has its own set of segments in common.
3. In some cases, all four combinations of pairwise homologies in two segments occur. For example, λ and 82 homoduplex in the 3′ end of the tail gene *J;* PA-2 and 21 are likewise homologous with each other, but not with λ or 82. On the other hand, in a segment including the *cIII* gene, λ matches 21, and 82 matches PA-2. This is the expected result of homologous recombination in the *redX-kil* segment (shared by all four phages) and is difficult to explain without invoking recombination.

Annual Reviews
www.annualreviews.org/aronline

4. Breakpoints between segments appear to be sharp. For example, phage 434 matches λ in the integrase operon, then becomes like PA-2 to the right. At the resolution of heteroduplex mapping (50–100 bp), no gap or overlap is apparent.
5. Breakpoints recur at some sites. λ, 82, and 21 all match in the *P/ren* segment, but diverge to three alternatives for the cluster of accessory genes to the right. Clearly at least two independent DNA-joining events must have occurred at this position.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

These conclusions have been refined by examination of the substantial sequence database now available for phages and defective prophages (25). The only lambdoid phage to be sequenced in its entirety is λ (49 kb—this is laboratory λ, not Ur-λ), but approximately 43 kb of P22, 16 kb of 21, 10 kb of φ80, and 9 kb of HK022 have been recorded by GENBANK. (These entries have not been thoroughly screened for overlaps.) All available sequence data corroborate the heteroduplex analysis in showing that each lambdoid genome is a patchwork of segments related to various other members of the group.

The sharpness of the breakpoints is illustrated by the integration genes of λ, 434, and HK022 (5). The leftmost (downstream) element of the integrase operon is the *sib* site. Within a few nucleotides downstream of *sib*, these three phages diverge completely from one another. Upstream of the integrase operon lies an open reading frame (ORF55) of unknown function, and the strong terminator *t*L4. Upstream of *t*L4, the three phages again diverge. Because the DNA downstream of *sib* and upstream of *t*L4 has no known function in λ biology, one might argue that these sequences have simply drifted in absence of selection pressure; however, the upstream sequence of HK022 is also found in the same location in phage 21, requiring a recombinational origin.

Whereas the divergence of λ from 434 is virtually complete upstream of *t*L4, 434 and HK022 are weakly homologous (55% sequence identity) for 80 bp upstream of *t*L4. This could reflect an earlier upstream crossover.

Within the integrase operon, λ and 434 match closely, except that 434 has an insertion relative to λ of 441 bp, including one open reading frame, between *xis* and ORF55. The same insertion is found in HK022. The DNA extending from *t*L4 downstream through the middle of *int* is closely homologous for these three phages (less than 5% divergence, with synonymous substitutions predominating in the open reading frames.) The 3′ ends of *int* and *attP* are likewise similar for λ and 434, but much more divergent in HK022, consistent with the functional divergence in core recognition. Homology is recognizable through *sib*, but not beyond. Thus a recombination in the middle of *int* may have occurred between a 434-like ancestor and a partner that contributed the downstream sequence. The alternative hypothesis that the downstream changes in HK022 came about through mutation plus selection for altered function is

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

unlikely because of the prevalence of synonymous codon changes. Within the rest of *int-xis* DNA, a few other possible exchange points can be identified.

These examples indicate that breakpoints between close homology and undetectable homology can be sharp, that they recur at specific places on the genome, and that homologous intragenic recombination has occurred between partners that had already diverged extensively (as in HK022 and λ *int* genes). Breakpoints can occur at natural boundaries between functional units; the *sib* sequence is the furthest downstream element of the integrase operon, and *t*L4 protects that operon from upstream transcription. The *sib* and *t*L4 sequences are conserved for functional reasons, and therefore could constitute microhomologous targets for recombination between otherwise heterologous phages. These sequences are exactly at the breakpoints; heterology starts immediately beyond them, rather than in their general vicinity, and except for HK022 vs 434 upstream of *t*L4, there is no intermediate zone of partial homology.

A computer search of GENBANK for contigs of other phages that match λ for better than 90% for more than 100 bp at some place and drop to less than 35% elsewhere can help us determine whether these features are general or can be dismissed as anecdotal (C Burge, personal communication): Phages 21 and PA-2 are very similar in the lysis genes and diverge from λ at the boundary between *R* (nonhomologous) and *Rz* (homologous). The homology starts with the initiator codon of *Rz* (12, 25). Thus the breakpoint is at a special position, but a conserved microhomology larger than ATGA is not evident. Comparison of the *Rz* sequences of P22, HK97, and λ also indicate two homologous intragenic crossovers for P22 vs λ (24, 25). The first 60 bp of P22 are 62% identical to λ; the second 60 are 92% identical; and the third are 52% identical. HK97 is like λ for the first 60 bp, like both λ and P22 for the next 60, then like P22 for the rest of the gene. Thus HK97 resembles a crossover between λ-like parent and a P22-like parent within the second 60 bp. Amino acid sequence conservation is higher in this segment (95–100% vs 70–75% on either side.) However, conservation of protein structure does not fully explain the DNA sequence similarity because (as in the case of HK022 *int* vs λ *int*), the frequency of synonymous substitutions is significantly lower within the conserved segment than upstream or downstream from it.

P22 lacks a functional *ren* gene but becomes homologous to λ downstream of *ren*, through the intergenic DNA and the downstream open reading frames in the *P-Q* interval (4). A 42-bp segment surrounding the *ren* terminator has intermediate homology (about 64%). Phage 434 has minimal homology with λ in *cI* and *cro*, then suddenly becomes homologous 5 bp downstream of *cro* (47). P22 is homologous to λ through *ral* and the intergenic DNA upstream of it, but diverges 16 bp before the start of the λ *sieB* (*git*) gene (38).

The segment of intermediate homology at the terminus of P22*ren* may indicate more than one crossover, as suggested earlier for HK022 and 434

Annual Reviews
www.annualreviews.org/aronline

upstream of *t*L4. Two other cases bear evidence for multiple events. The first (HK022 vs λ between *cIII* and *ral*) is probably pure coincidence. In this interval, HK022 has experienced an insertion of IS*903* and a close (but not adjacent) deletion of *ssb* relative to λ; 50 bp upstream of the IS*903* lies a breakpoint suggesting an interstrain recombination 8 bp downstream from HK022 *nun* (74). The second (phage 21 vs λ upstream of *Q*) is more intriguing (48). λ and 21 match in the upstream part of the *P-Q* interval, including ORF209 and the 5′ end of ORF68. In phage 21 the rest of ORF68 and the next downstream gene are missing, putting the other side of the breakpoint just upstream from *Q*. The *Q* genes of λ and 21 barely match. However, upstream of *Q* (fused to ORF68 in 21), both phages have the sequence TCANNAGGA-GAAAGGCGCATG, which includes the ribosome-binding site and initiation codon for *Q*. 21 could be the product of two separate events—deletion in a λ-like parent, preceded or followed by recombination at this microhomology; however, the fact that the deletion ends exactly at this position suggests a coordinated event, as did the joint occurrence of limited homology recombination and illegitimate recombination in the genesis of some λ *crg* isolates (80).

This small survey shows that the properties generalized from the integrase operon are neither atypical nor universal. Most breaks are sudden, the exceptions being HK022 vs 434 upstream of *t*L4 and P22 vs λ downstream of *ren*. Many lie at or near functional boundaries, although this is not true for P22 vs λ between *ral* and *sieB* (intergenic) or 21 and λ in ORF68 (intragenic).

Genesis of the New Combinations

Clearly, natural lambdoid phages are related by recombination. Those recombinants that have become established must have passed the test of natural selection. Three questions can now be asked: (*a*) What, if anything, can be inferred about the recombinational mechanisms used? (*b*) How far has natural selection restricted the products of recombination? (*c*) What are the consequences for the population structure of the lambdoid phages? The third question is deferred to the end. This section discusses the first two.

When two phage types encounter one another in nature, four types of recombination might occur (25): (*a*) Recombination within homologous segments can occur between any two lambdoid phages, which will be homologous in some segments and should occasionally recombine there. (*b*) Site-specific recombination (known only at the *att* site used in integration) sometimes transpires. (*c*) Recombination can happen at microhomologies of oligonucleotide length, such as *sib* or *t*L4. Such sites, whose sequences are conserved for functional reasons, may constitute the preferred exchange points for recombination between partners that diverge on both sides of the site. (*d*) Ille-

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

gitimate recombination can also occur at positions where the recombining partners have no homology whatever.

Of these four processes, the first two can proceed at high rates, although the rate of homologous recombination drops rapidly with the percent of mismatching between the partners (79). Thus the observed rate of recombination between natural phages can be lower than that between mutants of the same phage by several orders of magnitude (66); the intragenic recombinations inferred in *int* (HK022 vs λ) and perhaps in *Rz* are noteworthy for this reason. Homologous recombination can create new combinations of flanking genes, but it cannot create new breakpoints between homology and nonhomology. When compared with parent A, a given recombinant is identical to that parent to the left of the recombination point; to the right, it shows the same breakpoints that distinguish parent A from parent B. Compared with parent B, the recombinant shows parental breakpoints to the left of the recombination point. Neither does integrase-catalyzed site-specific recombination create new breakpoints. It occurs only between phages with the same integrase specificity, and in all known cases these are homologous on both sides of the *att* site.

The other two mechanisms can generate sharp breakpoints. For example, if two phages have diverged extensively in the DNA surrounding *t*L4 but have conserved the termination site itself, microhomologous recombination can produce a recombinant resembling parent A to the left of *t*L4 and parent B to the right. When the recombinant is tested against either parent, a new breakpoint appears that is not observed when the parents are compared with each other. When parents have diverged less extensively, recombination at *t*L4 may create a transition from close homology to partial homology (as seen with HK022 and 434 to the right of *t*L4). Microhomologous recombination will not by itself create a further decrease in homology with distance from the recombination point. That requires at least one additional recombination event (to give a zone of uniform partial homology) or multiple events (to give a gradual decrease in homology with distance). The known examples (HK022 vs *t*L4 and P22 vs λ at *ren*) do not distinguish a uniform zone from a gradual decrease. Illegitimate recombination also creates new breakpoints.

One observes breakpoints when comparing a recombinant to a single parent; the other parent or the reciprocal recombinant is rarely included in the collection compared. Thus, there is usually no way to decide which strain is the parent and which is the recombinant.

The first three mechanisms automatically generate recombinants with a full complement of phage genes (because the exchange points are the same on both partners). Illegitimate recombination occurs at the lowest rate, and the vast majority of its products have more or less than a complete genome. However, it is the only mechanism that creates unique breakpoints at arbitrary positions on heterologous partners.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Explanations for the existing array of lambdoid phages vary in the relative weight they accord to recombination and to the subsequent natural selection. At one extreme, the types we find are informative as to mechanisms of origin; although they must represent the small minority that had at least a local selective advantage, the most common events (especially looking in the immediate neighborhood of the recombination points) are likely to be represented most frequently. At the other extreme, natural selection is all-important. When illegitimate recombination is included, any gene combination can appear, and the natural collection of phages represents the endpoint of selective optimization. The truth doubtless lies in between, but the extremes help define the issues.

Susskind & Botstein (94) made the most specific suggestion putting the major burden for new combinations on recombination before λ had been fully sequenced. They proposed that functional genes are separated from one another by functionless recombinational linkers, the latter being conserved solely for the purpose of promoting recombination. Such a structure would allow frequent natural recombination with conservation of gene order, so that all the gene combinations constructible by mixing and matching would be available to natural selection. However, the evolutionary stability of such a structure is problematic, and DNA sequencing leaves little room for conserved functional spacers. Most closely resembling these spacers are microhomologies, such as those at *sib*, *tL4*, and the translational start site for *Q*, which are apparently conserved for function and should allow a low rate of exchange at corresponding points on the two partners, thus reducing the chance that one lambdoid phage might diverge so extensively that it falls outside of the lambdoid gene pool.

Of course, with a sufficient rate of illegitimate recombination, no lambdoid phage can fall outside the pool (which in fact can include any DNA with which the phages come in contact). The location and recurrence of breakpoints could be attributable to the power of natural selection. Breakpoint location might then be rather uninformative on their genesis, because multiple events could have culminated in selective optimization.

One fact that is most easily explained by simple microhomologous recombination is the location of sharp breakpoints immediately adjacent to *sib* and *tL4* (5). The distal DNA is intergenic, with no known function. Illegitimate recombination might have generated breakpoints 10–100 bp beyond the sites. No appreciable selective advantage to eliminating those 10–100 bp is evident, so it is easiest to think that the observed sequences record the primary event. Other cases of proximity to gene boundaries are less clearcut. For example, the length of DNA between *cro* and *nutR* is 26 bp. If selective constraints place the crossover between these two elements, the location of the λ/434 breakpoint 5 bp downstream of *cro* is unremarkable. And genes *R* and *Rz* are contiguous,

Annual Reviews
www.annualreviews.org/aronline

so perhaps the terminator-initiator ATGA is the only breakpoint that could generate a viable recombinant.

Molecular Phylogeny

Their extensive natural recombination prevents us from arranging lambdoid phage genomes on a family tree. Individual genes can generally be so arranged, although intragenic recombination may negate even that possibility as the database expands. The conspecific integrases of λ and 434 are almost identical; the closest heterospecific relative is HK022; and 21 integrase lies on a different branch (5, 84). However, immediately upstream of *t*L4, 21 and HK022 are closely related, but λ and 434 are unrelated to these phages or to each other. In this respect, individual natural isolates resemble members of a sexual species rather than asexually evolving lineages.

Selection for Polymorphism

Regardless of what recombinational mechanisms generate new types, natural selection must play a major role in maintaining a population with diverse gene combinations. The survival of the products of rare events such as microhomologous or illegitimate recombination (especially of recurrent microhomologous recombinations at the same site) argues that the gene combinations generated must have some special advantages.

A selective advantage of the genotype Ab arising from the cross ABXab is possible only if the two parents differ at both locus A and locus B; if numerous recombinations at different positions have produced advantaged recombinants, functional polymorphism must occur at many loci. Beyond the advantages of recombinants, how the polymorphisms themselves arise and persist merits attention. For the survival of diverse specificities in processes such as repression, integration, and termination, the simplest explanation is frequency-dependent selection (20).

The idea can be illustrated with repression. Evolutionary change from one specificity to another requires alteration of both repressor and Cro, as well as the two tripartite operators they control. Therefore, a considerable kinetic barrier makes origin by means of random drift unlikely. Imagine a universe in which all lambdoid phages had the same repression specificity as λ. Some *E. coli* cells would be lysogenic for λ and others sensitive to it, in an equilibrium dependent on the free phage concentration. Phage multiplication would be limited by the fact that only nonlysogens can be productively infected. Within this universe, imagine the (very rare) appearance of a new specificity, such as 434. The 434 phage will productively infect every cell it encounters and therefore can outcompete λ. The competitive advantage remains until 434 becomes as common as λ. Analogous arguments can be made for other specificities, such as integrase.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

How do polymorphisms arise? In some cases, such as repression and integration, the homologies among different lambdoid phages suggest mutational divergence in situ from a primordial ancestral phage. Occasionally, genes or gene segments appear to have been introduced from sources outside the lambdoid family. These include parts of some integrase genes. Site-specific recombinases related to λ integrase are widely distributed among both prokaryotes and eukaryotes, and the C-terminal ends of the integrases of P22 and φ80 are related to the C termini of other lambdoid phages (3). As all the integrases are apparently of common ancestry, these may represent ancient recombinational events similar to that relating HK022 to 434. Other examples are the substitution in phage P22 of a *dnaB* homologue in place of λ *P* (13) and the replacement of λ *R* with lysozyme genes in other lambdoid phages (12). The tail-fiber genes also have homologies to those of nonlambdoid phages; in fact, the sequences of tail-fiber genes suggest that numerous such intergroup transfers have survived, presumably selected for infectivity on locally available hosts (25, 49). Thus, genes have been pirated on occasion from various sources, with no indication of flanking homology. The fact that such replacements of analogues (perhaps by illegitimate recombination) preserve gene order suggests a strong selection for conservation of order to preserve regulatory circuitry.

Where specificity changes have arisen by mutational divergence from a common ancestor (as appears to be the case for lambdoid repressor genes, for instance) the observed sequence divergence far exceeds the minimal alterations needed to change specificity (101). The divergence suggests that the different specificity types are relatively ancient. Some of the differences may have been selected for fine tuning following the primary change, whereas others may have come about purely by drift.

Accessory Genes

In one view of lambdoid phage evolution, the common ancestor of the group contained only the basic elements essential to the productive and lysogenic cycles; genes for luxury functions were added later, some of them quite recently. Once present in one lambdoid phage, a gene that conferred some modest advantage would spread through the gene pool. Facts that might favor this viewpoint include the irregular distribution of such genes among lambdoid phages and the absence of extensive sequence polymorphism.

For example, a gene (*nmpC*, for new membrane protein), related to the bacterial outer membrane protein gene *ompC* and presumably pirated from the bacterial chromosome, is present to the right of *Rz* in phage PA-2 and in defective prophage DLP12 (54, 67); *nmpC* is not present in λ nor most of the well-studied lambdoid phages, but it can be found in other natural isolates screened for this property. Phage PA-2 uses *ompC* as a receptor in attachment

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

(as can Ur-λ). NmpC synthesis turns off OmpC synthesis, rendering the cell surface unable to adsorb PA-2.

The fact that *nmpC* is not closely linked to the tail-fiber genes might indicate that it is a relatively recent addition that has not yet found its optimal location. However, if it was acquired recently and is spreading through the lambdoid population by homologous recombination, the flanking DNA in PA-2 and DLP12 should be homologous. It isn't, although the location in the genome is similar for PA-2 and DLP12. Similarly, the sequence of the restriction alleviation gene *ral* is identical in phages λ and 21 (with no flanking homology) though absent from the same location in phage 82 (18, 38).

The *rexAB* genes (which interfere with multiplication of some other phages) are present in λ (downstream of and contranscribed with *cI*, Figure 1) but lacking in the other common lambdoid phages. The mechanism of exclusion seems to be a collapse of the membrane potential triggered by infection of a λ lysogen by a second phage, killing the cell and preventing phage development (76). λ has a second gene, *ren* (downstream from *P*), which is present in 434, 21, and φ82 but absent from P22 and 80, that renders lambdoid phage replication Rex-resistant. Both RexA and RexB are needed, in the right proportions, for the exclusion; excess RexB neutralizes their effects. Excess *RexB* can also stabilize gp*O* λ expressed from a plasmid (86). The *rexB* gene can be transcribed from a second promoter, so its in vivo expression need not always be coordinate with *rexA*.

Despite the pleiotropic effects of RexB, the simplest interpretation is that the natural function of RexAB is to exclude other phages and that the natural function of Ren is to resist exclusion. The natural exclusion system targeted by Ren need not be Rex. In fact, Ren is not needed when lambdoid phages infect wild-type λ lysogens; it is needed only in the unnatural circumstance of infecting lysogens of λ *cIts* mutants that overproduce Rex (97). The absence of *ren* in some lambdoid phages could mean that it is a recent addition to the genome, but an equally plausible explanation is that it was deleted from some phages not exposed to strong selection.

SPECIATION IN PHAGES

The lambdoid phages may collectively be regarded as a biological species with its own gene pool (18). This supposition has two implications: frequent recombination within the species and barriers to introgression of DNA from other sources.

The first criterion is clearly met. Natural recombination may take place in coinfected cells, in infected lysogens, or in infected defective lysogens. The third is probably the most common confrontation, because defective prophages are present in most natural cell lines. Thus the defective prophages are signifi-

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

cant components of the gene pool and cannot be dismissed as evolutionary dead ends. As most defective prophages contain limited segments of phage genomes interrupted by deletions or insertions, the typical recombination events are expected to be double crossovers exchanging short DNA segments of the infecting phage with their homologues.

Given that most members of the gene pool are related by dispersed homologies, exchange by homologous recombination is likely to be the most frequent process generating new combinations. This idea underlies the linker hypothesis and has dominated most subsequent discussions. However, the most frequent process need not be the most important. Whereas homologous recombination is frequent within the gene pool, illegitimate recombination recognizes no such boundaries and may have recruited any type of heterologous DNA with which the phage comes in contact.

The fact that some lambdoid phages have appropriated genes from host chromosomes or nonlambdoid phages may seem to undermine the applicability of the species concept. However, the concept is inherent in the language most workers use. Thus, the acquisitions of foreign DNA are usually classified as lateral transfers, a term that presupposes a qualitative difference from recombination within members of the same group. The notion that the group constitutes a distinct gene pool is implicitly recognized more often than acknowledged.

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Acknowledgments

The author thanks Chris Burge for information from a computer search of GENBANK and several colleagues who supplied reprints and manuscripts. The author's research is supported by a grant AI08573, National Institute of Allergy and Infectious Diseases.

Any *Annual Review* chapter, as well as any article cited in an *Annual Review* chapter, may be purchased from the Annual Reviews Preprints and Reprints service. 1-800-347-8007; 415-259-5017; email: arpr@class.org

Literature Cited

1. Aggarwal A, Rodgers DW, Drottar M, Ptashne M, Harrison SC. 1988. Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science* 242:899–907
2. Anilionis A, Riley M. 1980. Conservation and variation of nucleotide sequences within related bacterial genomes: *Escherichia coli* strains. *J. Bacteriol.* 143:355–65
3. Argos P, Landy A, Abremski K, Egan JB, Haggard-Ljungquist E, et al. 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J.* 5:433–40
4. Backhaus H, Petri JB. 1984. Sequence analysis of a region from the early right operon in phage P22 including the replication genes 18 and 12. *Gene* 32:289–303
5. Baker J, Limberger R, Schneider SJ,

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Campbell A. 1991. Recombination and modular exchange in the genesis of new lambdoid phages. *New Biol.* 3: 297–308
6. Balakin AG, Skriphin EA, Shatsky IN, Bogdanov AA. 1992. Unusual ribosome binding properties of mRNA encoding bacteriophage λ repressor. *Nucleic Acids Res.* 20:567–71
7. Baron J, Weisberg R. 1992. Mutations of the phage λ *nutL* region that prevent the action of Nun, a site-specific transcription termination factor. *J. Bacteriol.* 174:1983–89
8. Barondess JJ, Beckwith J. 1990. A bacterial virulence determinant encoded by lysogenic coliphage lambda. *Nature* 346:871–74
9. Beamer LJ, Pabo CO. 1992. Refined 1.8 Å crystal structure of the λ repressor-operator complex. *J. Mol. Biol.* 227: 177–96
10. Béjar S, Bouché F, Bouché J-P. 1988. Cell division inhibition gene *dicB* is regulated by a locus similar to lambdoid bacteriophage immunity loci. *Mol. Gen. Genet.* 212:11–19
11. Bezdek M, Amati P. 1967. Properties of P22 and a related *Salmonella typhimurium* phage. *Virology* 31:212–18
12. Bonovich MT, Young R. 1991. Dual start motif in two lambdoid S genes unrelated to λ S. *J. Bacteriol.* 173:2897–2905
13. Botstein D, Chan RK, Waddell CH. 1972. Genetics of bacteriophage P22. II. Gene order and gene function. *Virology* 49:268–82
14. Bouché J-P, Béjar S, Cam K. 1990. Cooption of prophage genes: new data on the Kim prophage region of the *Escherichia coli* chromosome. See Ref. 32a, pp. 373–80
15. Brody H, Hill CW. 1988. Attachment site of the genetic element e14. *J. Bacteriol.* 170:2040–44
16. Burland V, Plunkett G III, Daniels DL, Blattner FR. 1993. DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome. Organizational symmetry around the origin of replication. *Genomics* 16:551–61
17. Bushman FD. 1993. The bacteriophage 434 right operator. Roles of O(R)1, O(R)2 and O(R)3. *J. Mol. Biol.* 230:28–40
18. Campbell A. 1988. Phage evolution and speciation. In *The Bacteriophages,* ed. R Calendar, pp. 1–14. New York: Plenum. 596 pp.
19. Campbell A. 1992. Chromosomal insertion sites for phages and plasmids. *J. Bacteriol.* 174:7495–99
20. Campbell A, Botstein D. 1983. Evolution of the lambdoid phages. See Ref. 52a, pp. 365–80
21. Campbell A, Schneider SJ, Song B. 1992. Lambdoid phages as elements of bacterial genomes. *Genetica* 86:259–67
22. Carlson NG, Little JW. 1993. Highly cooperative DNA binding by the coliphage HK022 repressor. *J. Mol. Biol.* 230:1108–30
23. Carlson NG, Little JW. 1993. A novel anti-virulence element in the temperate bacteriophage HK022. *J. Bacteriol.* 175:7541–49
24. Casjens S, Eppler K, Parr R, Poteete AR. 1989. Nucleotide sequence of the bacteriophage P22 gene 13 to 3 region. Identification of a new gene required for lysis. *Virology* 171:588–98
25. Casjens S, Hatfull G, Hendrix R. 1992. Evolution of ds DNA tailed-bacteriophage genomes. *Sem. Virol.* 3:383–97
26. Court D, Oppenheim AB. 1983. Phage lambda's accessory genes. See Ref. 52a, pp.251–77
27. Craig NL. 1988. The mechanism of conservative site-specific recombination. *Annu. Rev. Genet.* 22:77–106
28. Cue D, Feiss M. 1992. Genetic analysis of mutations affecting terminase, the bacteriophage λ DNA packaging enzyme, that suppress mutations in *cosB,* the terminase binding site. *J. Mol. Biol.* 228:72–87
29. Cue D, Feiss M. 1993. A site required for termination of packaging of the phage λ chromosome. *Proc. Natl. Acad. Sci. USA* 90:9290–94
30. Das A. 1992. How the phage lambda N gene product suppresses transcription termination: communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA. *J. Bacteriol.* 174:6711–16
31. Dhillon EKS, Dhillon TS, Lam YY, Tsang AHC. 1980. Temperate coliphages: classification and correlation with habitats. *Appl. Environ. Microbiol.* 39:1049–53
32. Doelling JH, Franklin NC. 1989. Effect of all single base substitutions in the loop of *boxB* on antitermination of transcription by bacteriophage λ's N protein. *Nucleic Acids Res.* 17:5565–77
32a. Drlica K, Riley M, eds. 1990. *The Bacterial Chromosome.* Washington DC: Am. Soc. Microbiol. 469 pp.
33. Dubose RF, Dykhuizen DE, Hartl DL. 1988. Genetic exchange among natural isolates of bacteria: recombination within the *phoA* gene of *Escherichia coli. Proc. Natl. Acad. Sci. USA* 85: 7036–40
34. Echols H, Guarneros G. 1983. Control

Annual Reviews
www.annualreviews.org/aronline

of integration and excision. See Ref. 52a, pp. 251–77
35. Espion D, Kaiser, Dambly-Chaudiere C. 1983. A third defective lambdoid prophage of *Escherichia coli* K12 defined by the λ derivative λ *qin111*. *J. Mol. Biol.* 170:611–33
36. Feiss M, Sippy J, Miller G. 1985. Processive action of terminase during sequential packaging of bacteriophage λ chromosomes. *J. Mol. Biol.* 186:759–71
37. Fisher R, Feiss M. 1980. Reversion of a cohesive end site mutant of bacteriophage lambda by recombination with a defective prophage. *Virology* 107:160–73
38. Franklin NC. 1985. Conservation of genome form but not sequence in the transcription antitermination determinants of bacteriophages λ, ϕ21, and P22. *J. Mol. Biol.* 181:75–84
39. Franklin NC. 1992. The carboxy-terminal 14 amino acids of phage λ N protein are dispensable for transcription antitermination. *J. Bacteriol.* 174:8144–47
40. Friedman DI, Cranston AE, Thompson D, Schauer AT, Olsen ER. 1989. Genetic analysis of the N transcription antitermination system of phage λ. *Genome* 31:491–96
41. Furth M, Yates J, Dove W. 1979. Positive and negative control of DNA replication. *Cold Spring Harbor Symp. Quant. Biol.* 43:147–53
42. Gillen JR, Karu AE, Nagaishi H, Clark AJ. 1977. Characterization of the deoxyribonuclease determined by lambda reverse as exonuclease VIII of *Escherichia coli*. *J. Mol. Biol.* 113:27–41
43. Goliger JA, Roberts JW. 1989. Sequences required for antitermination by phage 82 Q protein. *J. Mol. Biol.* 210:461–71
44. Gottesman MM, Gottesman ME, Gottesman S, Gellert M. 1974. Characterization of bacteriophage lambda reverse as an *Escherichia coli* phage carrying a unique set of host-derived recombination functions. *J. Mol. Biol.* 88:471–87
45. Greener A, Hill CW. 1980. Identification of a novel genetic element in *Escherichia coli* K-12. *J. Bacteriol.* 144. 312–21
46. Grodberg J, Lundrigan MD, Toledo DL, Mangel WF, Dunn JJ. 1988. Complete nucleotide sequence and deduced amino acid sequence of the *ompT* gene of *Escherichia coli* K-12. *Nucleic Acids Res.* 16:1209
47. Grosschedl R, Schwarz E. 1979. Nucleotide sequence of the *cro-cII-oop* region of bacteriophage 434 DNA. *Nucleic Acids Res.* 6:867–81
48. Guo H-C, Kainz M, Roberts JW. 1991. Characterization of the late-gene regulatory region of phage 21. *J. Bacteriol.* 173:1554–60
49. Haggard-Ljungquist E, Halling C, Calendar R. 1992. DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J. Bacteriol.* 174:1462–77
50. Hase T, Nakai M, Masamure Y. 1989. Transcription of a region downstream from λ *ori* is required for replication of plasmids derived from coliphage lambda. *Mol. Gen. Genet.* 216:120–25
51. Hauser MA, Scocca JJ. 1992. Site-specific integration of the *Haemophilus influenzae* bacteriophage HP1: location of the boundaries of the phage attachment site. *J. Bacteriol.* 174:6674–77
52. Hendrix RW, Duda RL. 1992. λ PaPa is not the mother of all λ phages. *Science* 258:1154–58
52a. Hendrix RW, Roberts JW, Stahl FW, Weisberg RA, eds. 1983. *Lambda II*. Cold Spring Harbor, NY: Cold Spring Harbor Lab. 694 pp.
53. Hershey AD, Dove W. 1971. Introduction to lambda. See Ref. 53a, pp. 3–11
53a. Hershey AD, ed. 1971. *The Bacteriophage λ*. Cold Spring Harbor NY: Cold Spring Harbor Lab. 792 pp.
54. Highton PJ, Chang Y, Marcotte WR, Schnaitman CA. 1985. Evidence that the outer membrane protein gene *nmpC* of *Escherichia coli* K-12 lies within the defective qsr′ prophage. *J. Bacteriol.* 162:256–62
55. Highton PJ, Chang Y, Myers R. 1990. Evidence for the exchange of segments between genomes during the evolution of lambdoid bacteriophages. *Mol. Microbiol.* 4:1329–40
56. Ho YS, Pfarr D, Strickler J, Rosenberg M. 1992. Characterization of the transcription activator protein cI of bacteriophage P22. *J. Biol. Chem.* 267: 14388–97
57. Jacob F, Campbell A. 1959. Sur le systéme de repression assurant l'immunité chez les bacteries lysogenes. *C. R. Acad. Sci.* 248:3219–21
58. Johnson G, Widner W, Xin W, Feiss M. 1991. Interference with phage λ development by the small subunit of the phage 21 terminase, gp1. *J. Bacteriol.* 173:2733–38
59. Kaiser AD. 1957. Mutations in a temperate bacteriophage affecting its ability to lysogenize. *Virology* 3:42–61
60. Kaiser AD, Jacob F. 1957. Recombina-

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Reviews
www.annualreviews.org/aronline

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

tion between related temperate bacteriophages and the genetic control of immunity and prophage localization. *Virology* 4:509–21

61. Kaiser K, Murray NE. 1980. On the nature of *sbc* mutations in *E. coli* K-12. *Mol. Gen. Genet.* 179:555–63
62. Kim S, Landy A. 1992. Lambda Int protein bridges between higher order complexes at two distant chromosomal loci *attL* and *attR*. *Science* 256:198–203
63. Kitamura J, Mise K. 1970. A new generalized transducing phage in *Salmonella*. *Jpn. J. Med. Sci. Biol.* 23:99–102
64. Krylov VN, Tsygankov YD. 1976. New lambdoid phages of *Escherichia coli*. I. Isolation, immunity groups, and recombination with phage λ. *Soviet Genet.* 12:1123–29
65. Lazinski D, Grzadzielska E, Das A. 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* 57:207–18
66. Liedke-Kulke M, Kaiser AD. 1967. Genetic control of prophage insertion specificity in bacteriophages lambda and 21. *Virology* 32:465–74
67. Lindsey DR, Mullin DA, Walker JR. 1989. Characterization of the cryptic lambdoid prophage DLP12 of *Escherichia coli* and overlap of the DLP12 integrase gene with the tRNA gene *argU*. *J. Bacteriol.* 171:6197–6205
68. Low KB. 1973. Restoration by the *rac* locus of recombinant forming ability in *recB recC* merozygotes of *Escherichia coli* K12. *Mol. Gen. Genet.* 122:119–30
69. Luzzati D. 1970. Regulation of λ exonuclease synthesis: role of the N gene product and λ repressor. *J. Mol. Biol.* 49:515–19
70. Mason SW, Li J, Greenblatt J. 1992. Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. *J. Biol. Chem.* 267:19418–26
71. Nagaraja R, Weisberg RA. 1990. Specificity determinants in the attachment sites of bacteriophages HK022 and λ. *J. Bacteriol.* 172:6540–50
72. Nunes-Duby SE, Matsumoto L, Landy A. 1987. Site-specific intermediates trapped with suicide substrates. *Cell* 50:779–88
73. Oberto J, Clerget M, Ditto M, Cam K, Weisberg RA. 1993. Antitermination of early transcription in phage HK022. Absence of a phage-coded antitermination factor. *J. Mol. Biol.* 229:368–81
74. Oberto J, Weisberg RA, Gottesman ME. 1989. Structure and function of the *nun* gene and the immunity region of the lambdoid phage HK022. *J. Mol. Biol.* 207:675–93
75. Ordal GW, 1971. Supervirulent mutants and the structure of operator and promoter. See Ref. 53a, pp. 565–70
76. Parma DH, Snyder M, Sobolevski S, Mowray M, Brody E, Gold L. 1992. The Rex system of bacteriophage λ: tolerance and altruistic cell death. *Genes Dev.* 6:497–510
77. Patterson TA, Zhang Z, Baker T, Johnson, LL Friedman D, Court DL. 1993. Bacteriophage lambda N-dependent transcription antitermination: regulation by competition for an RNA site. *J. Mol. Biol.* Submitted
78. Ptashne M. 1986. *A Genetic Switch*. Cambridge, MA: Cell/Blackwell. 128 pp.
79. Rayssiguier C, Thaler DS, Radman M. 1989. The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* 342:396–401
80. Redfield RJ, Campbell A. 1987. Structure of cryptic prophages. *J. Mol. Biol.* 198:393–404
81. Riley M, Sanderson KE. 1990. Comparative genetics of *Escherichia coli* and *Salmonella typhimurium*. See Ref. 32a, pp. 85–95
82. Robledo R, Gottesman ME, Weisberg RA, 1990. λ *nutR* mutations convert HK022 Nun protein from a transcription termination factor to a suppressor of termination. *J. Mol. Biol.* 212:635–43
83. Schnaitman C, Smith D Forn de Salas M. 1975. Temperate bacteriophage which causes the production of a new major outer membrane protein by *Escherichia coli*. *J. Bacteriol.* 15:1121–36
84. Schneider SJ. 1992. *Site-specific recombination of lambdoid phage 21 into the* icd *gene of* Escherichia coli. PhD thesis. Stanford Univ., Stanford, CA. 293 pp.
85. Schnos M, Inman RB. 1991. Daughter origins can be held together by O protein in phage lambda replicative intermediates. *Virology* 183:753–56
86. Schoulaker-Schwarz R, Dekel-Gorodetsky L, Engelberg-Kulka H. 1991. An additional function for bacteriophage λ *rex*: the *rexB* product prevents degradation of λ O protein. *Proc. Natl. Acad. Sci. USA* 88:4996–5000
87. Senear DF, Lane TM, Ross JB, Waxman E, Eaton S, Rusinova E. 1993. The primary self-assembly reaction of bacteriophage lambda cI repressor dimers is to octamer. *Biochemistry* 32:6179–89
88. Shean CS, Gottesman ME. 1992. Trans-

Annual Reviews
www.annualreviews.org/aronline

lation of the prophage λ cI transcript. *Cell* 70:513–22
89. Smith MP, Feiss M. 1993. Sequence analyses of the phage 21 genes for prohead assembly and head completion. *Gene* 126:1–7
90. Smith MP, Feiss M. 1993. Sites and gene products involved in lambdoid phage DNA packaging. *J. Bacteriol.* 175:2393–99
91. Spanova A. 1992. Comparison of permuted region lengths in the genomes of related *Salmonella typhimurium* phages P22 and L. *Folia Microbiol.* 37:188–92
92. Strathern A, Herskowitz I. 1975. Defective prophage in *Escherichia coli* K-12 strains. *Virology* 67:136–48
93. Stoltzfus AB. 1991. *A survey of natural variation in the* trp-ton *region of the* E. coli *chromosome.* PhD. thesis, Univ. Iowa, Iowa City, IA. 234 pp.
94. Susskind MM, Botstein D. 1978. Molecular genetics of bacteriophage P22. *Microbiol. Rev.* 42:385–413
95. Taylor A. 1971. Endopeptidase activity of phage lambda-endolysin. *Nature New Biol.* 234:144–45
96. Thompson JF, Moitoso de Vargas L, Koch C, Kahmann R, Landy A. 1987. Cellular factors couple recombination with growth phase: characterization of a new component in the λ site-specific recombination pathway. *Cell* 50:901–8
97. Toothman P, Herskowitz I. 1980. Rex-dependent exclusion of lambdoid phages. II. Determinants of sensitivity to exclusion. *Virology* 102:147–60
98. Uetake H, Luria SE, Burrous JW. 1958. Conversion of somatic antigens in *Salmonella* by phage infection leading to lysis or lysogeny. *Virology* 5:68–91
99. Wegrzyn G, Pavlowicz A, Taylor K. 1992. Stability of coliphage λ DNA replication initiator, the λ O protein. *J. Mol. Biol.* 226:675–80
100. Weisberg R, Landy A. 1983. Site-specific recombination in phage lambda. See Ref. 52a, pp. 211–50
101. Wharton RP, Brown EL, Ptashne M. 1984. Substituting an α-helix switches the sequence specific DNA interactions of a repressor. *Cell* 38:361–69
102. Yagil E, Dolev S, Oberto J, Kislov N, Ramaiah N, Weisberg RA. 1989. Determinants of site-specific recombination in the lambdoid coliphage HK022. An evolutionary change in specificity. *J. Mol. Biol.* 207:695–717
103. Yang Y, Golliger JA, Roberts JW. 1989. Specificity and mechanism of antitermination by Q proteins of bacteriophage λ and 82. *J. Mol. Biol.* 210:453–60
104. Young R. 1992. Bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.* 56:430–81
105. Zylicz M, Arg D, Liberek K, Georgopolous C. 1987. Initiation of λ DNA replication with purified host- and bacteriophage-encoded proteins: the role of the *dnaK*, *dnaJ* and *dnaE* heat shock proteins. *EMBO J.* 8:1601–8

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org
by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

Annual Review of Microbiology
Volume 48, 1994

CONTENTS

A CHARMED LIFE, *Boris Magasanik* 1

GENETIC CONTROLS FOR THE EXPRESSION OF SURFACE ANTIGENS IN AFRICAN TRYPANOSOMES, *Etienne Pays, Luc Vanhamme, and Magali Berberof* 25

THE ROLE OF THE SIGMA FACTOR σ^S (KATF) IN BACTERIAL GLOBAL REGULATION, *Peter C. Loewen and Regine Hengge-Aronis* 53

THE EXTRACHROMOSOMAL DNAS OF APICOMPLEXAN PARASITES, *Jean E. Feagin* 81

TARGETING PROTEINS TO THE GLYCOSOMES OF AFRICAN TRYPANOSOMES, *Jürg M. Sommer and C. C. Wang* 105

HUMAN HIGH DENSITY LIPOPROTEIN KILLING OF AFRICAN TRYPANOSOMES, *S. L. Hajduk, K. M. Hager, and J. D. Esko* 139

GENETICS AND BIOCHEMISTRY OF DEHALOGENATING ENZYMES, *Dick B. Janssen, Frens Pries, and Jan R. van der Ploeg* 163

COMPARATIVE MOLECULAR BIOLOGY OF LAMBDOID PHAGES, *A. Campbell* 193

PATHWAYS AND MECHANISMS IN THE BIOGENESIS OF NOVEL DEOXYSUGARS BY BACTERIA, *Hung-wen Liu and Jon S. Thorson* 223

THE MOLECULAR PHYLOGENY AND SYSTEMATICS OF THE ACTINOMYCETES, *T. M. Embley and E. Stackebrandt* 257

LIVING BIOSENSORS FOR THE MANAGEMENT AND MANIPULATION OF MICROBIAL CONSORTIA, *Robert S. Burlage and Chiung-Tai Kuo* 291

IRON AND MANGANESE IN ANAEROBIC RESPIRATION: Environmental Significance, Physiology, and Regulation, *Kenneth H. Nealson and Daad Saffarini* 311

MECHANISMS UNDERLYING EXPRESSION OF TN*10*-ENCODED TETRACYCLINE RESISTANCE, *Wolfgang Hillen and Christian Berens* 345

BIOCHEMISTRY OF THE SOLUBLE METHANE MONOOXYGENASE, *John D. Lipscomb* 371

RAPID DETECTION OF FOOD-BORNE PATHOGENIC BACTERIA, *Bala Swaminathan and Peter Feng* 401

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.

HUMAN PAPILLOMAVIRUSES, *Harald zur Hausen and Ethel-Michele de Villiers* 427

THE ROLE OF pH AND TEMPERATURE IN THE DEVELOPMENT OF *LEISHMANIA* PARASITES, *Dan Zilberstein and Michal Shapira* 449

ANTIBIOTICS THAT INHIBIT FUNGAL CELL WALL DEVELOPMENT, *Manuel Debono and Robert S. Gordee* 471

STRUCTURAL AND FUNCTIONAL PROPERTIES OF *TRYPANOSOMA TRANS*-SIALIDASE, *S. Schenkman, D. Eichinger, M. E. A. Pereira, and V. Nussenzweig* 499

DESIGNING MICROORGANISMS FOR THE TREATMENT OF TOXIC WASTES, *Kenneth N. Timmis, Robert J. Steffan, and Ronald Unterman* 525

STRATEGIES FOR THE DISCOVERY OF SECONDARY METABOLITES FROM MARINE BACTERIA: Ecological Perspectives, *Paul R. Jensen and William Fenical* 559

MSCRAMM-MEDIATED ADHERENCE OF MICROORGANISMS TO HOST TISSUES, *Joseph M. Patti, Bradley L. Allen, Martin J. McGavin, and Magnus Höök* 585

COMPUTATIONAL DNA SEQUENCE ANALYSIS, *Samuel Karlin and Lon R. Cardon* 619

BIOLOGY AND GENETICS OF PRION DISEASES, *Stanley B. Prusiner* 655

AIDS-ASSOCIATED MYCOPLASMAS, *A. Blanchard and L. Montagnier* 687

ANTISENSE RNA CONTROL IN BACTERIA, PHAGES, AND PLASMIDS, *E. Gerhart H. Wagner and Robert W. Simons* 713

MICROBIAL IRON TRANSPORT, *Mary Lou Guerinot* 743

MICROBIAL BIOCATALYSIS IN THE GENERATION OF FLAVOR AND FRAGRANCE CHEMICALS, *Scott Hagedorn and Bryan Kaphammer* 773

INDEXES

Subject Index 801

Cumulative Index of Contributing Authors, Volumes 44–48 825

Cumulative Index of Chapter Titles, Volumes 44–48 827

Annu. Rev. Microbiol. 1994.48:193-222. Downloaded from arjournals.annualreviews.org by SUNY-Buffalo Health Scien on 02/21/05. For personal use only.