Supplementary Table 1A

Previous phage 21 sequence determinations

Accession No.	bp	location	number of differences from 21 sequence
M81255	8134	head genes	5
M23775	1688	head genes	5
AJ237660	4890	*cro-nin*	14
DQ372054	625	*cI-cII*	0
M58702	995	*Q*	0
M65239	1285	lysis	24
AH001308	191	early left operator	1
AH007390	949	*N*	29
M61865	2910	*int-att*	0
AF017628	1349	*int*	20
EU078592	1909	*N-cI-cro*	0

Supplementary Table 1B

Previous phage 434 sequence determinations

Accession No.	bp	location	number of differences from 434 sequence
M12904	970	*cI* region	4
Y00118	286	*cI* fragment	0
M12803	67	early left operator	0
X73093	116	early right operator	0
J02460	360	*cro*, early right operator	0
V00635	873	*cII-oop*	0
M60848	2616	*int*	0

Supplementary Table 2A

Phage 21 patches of genome similarity to λ

Current analysis

Coordinates	Percent identity	Simon & Davis coordinates[a]
1-8030	60.4%	1-8000[b]
8030-18130	95.7%	8000-18100
19220-19620	82.5%	[c]
29680-33660	97.5%	29200-33200
35520-38020	97.9%	35100-37700
38020-38990	97.6%	37800-38700
40590-41335	97.9%	40600-41500

Supplementary Table 2B

Phage 434 patches of genome similarity to λ

Current analysis

Coordinates	Percent identity	Simon & Davis coordinates[a]
1-18455	98.1%	1-17950
19508-20001	91.5	19100-19600
24826-26401	95.1%	24600-26500
26842-27088	95.2%	~2700 (short)
28061-31744	97.6%	28000-31700
32712-33200	92.5%	32900-33300
33968-34006	89.7%	34100-34100
34685-37522	97.9%	34800-37700
38056-39715	97.4%[d]	38300-39900
46454-47207	97.5%	46600-47500
47595-47993	96.0%	47700-47993

Table S2 footnotes

a. Calculated to three significant figures from the data reported in Simon *et al.* (1971), adjusting for their measurements of the 434 and 21 chromosomes to be about 48050 and 43400 bp long, respectively.

b. The region from 1-8000 gave variable heteroduplex in the presence of 40% formamide but heteroduplex was never observed in 60% formamide.

c. Short interval of moderately high or high similarity that was not observed in the reported heteroduplex analysis.

d. Value ignores a 185 bp insertion in λ DNA in this interval.

Supplementary Table S3A

Point differences between 434 and 434B

Gene	Coordinate (bp)	434 Nt	434B Nt	434 AA	434B AA	Codon
C	3970	A	G	Q	R	379
I	15193	T	C	S	P	108
J	18652	G	A	D	N	1050
J	18847	C	T	R	W	1116
cI	34178	A	C	L	R	61

Supplementary Table S3B

Point differences between λ DNA and λ portion of λ *imm434*

Gene	Coordinate (λ bp)	λ Nt	λ *imm434* Nt	λ AA	λ *imm434* AA	Codon
Left of *Nu1*	138	G	Δ[a]	–	–	–
K[c]	14266-7	Δ	G[a]	–	–	–
stf /*orf401*	20661	A	G[a]	K	E	338
tfa / *orf194*	22444	A	G[a]	K	E	158
ea59	25662	T	C[a]	M	V	438
orf63	31016	T	C[a]	N	D	61
exo	31966	A	T[a]	D	E	21
sieB	34934	A	G[a]	G	G	150
R	45618	T	C[a]	F	F	42

Supplementary Table S3C

Point differences between λ DNA and λ portion of λ *imm21*

Gene	Coordinate (λ bp)	λ Nt	λ *imm21* Nt	λ AA	λ *imm21* AA	Codon
Left of *Nu1*	138	G	Δ[a]	–	–	–
A	1864-1875	12 bp	Δ	EHYS	Δ	383-386
K[c]	14266-7	Δ	G[a]	–	–	–
J	17183	A	G	E	G	560
stf /*orf401*	19805	T	Δ[b]	D	fs	52
stf /*orf401*	20661	A	G[a]	K	E	338
stf /*orf401*	20832-3	Δ	C[b]	P	fs	396
stf /*orf401*	22444	A	G[a]	K	E	158
ea59	25662	T	C[a]	M	V	438
orf63	31016	T	C[a]	N	D	61
exo	31966	A	T[a]	D	E	21
R	45618	T	C[a]	F	F	42

Supplementary Table S3D

Point differences between 434 DNA and 434 portion of λ *h434 imm21*

Gene	Coordinate (434 bp)	434 Nt	λ *imm21 h4343* Nt	434 AA	λ *imm21 h434* AA	Codon
J	18651	G	A	D	N	1050
J	18847	C	T	R	W	1114
nleB	21678	G	A	V	I	5

Supplementary Table S3E

Point differences between λ DNA and λ portion of λ *h434 imm21*

Gene	Coordinate (λ bp)	λ Nt	λ *imm21 h4343* Nt	λ AA	λ *imm21 h434* AA	Codon
Left of *Nu1*	138	G	Δ[a]	–	–	–
K[c]	14266-7	Δ	G[a]	–	–	–
orf63	31016	T	C[a]	N	D	61
exo	31966	A	T[a]	D	E	21
R	45618	T	C[a]	F	F	42

Table S3 footnotes

a. These differences from the Sanger *et al*. (1982) λ sequence (Genbank Accession No. J02459) are present in all three of the hybrid phage genome sequences, except for the λ*imm434* difference at 34934 which is replaced by 21 DNA in the λ*imm21* hybrids.

b. Curiously, the λ *imm21 stf* gene does not carry the frameshift (deletion of one C) that is present at bp 20833 in most laboratory λ strains (Hendrix & Duda 1992), but it has a different frameshift at bp 19805 (λ coordinates).

c. The gene *K* start is incorrectly annotated in Accession No. J02459, and this is a frameshift error in that sequence (A. Davidson, personal communication).

Figure S1

Phage 21 genome

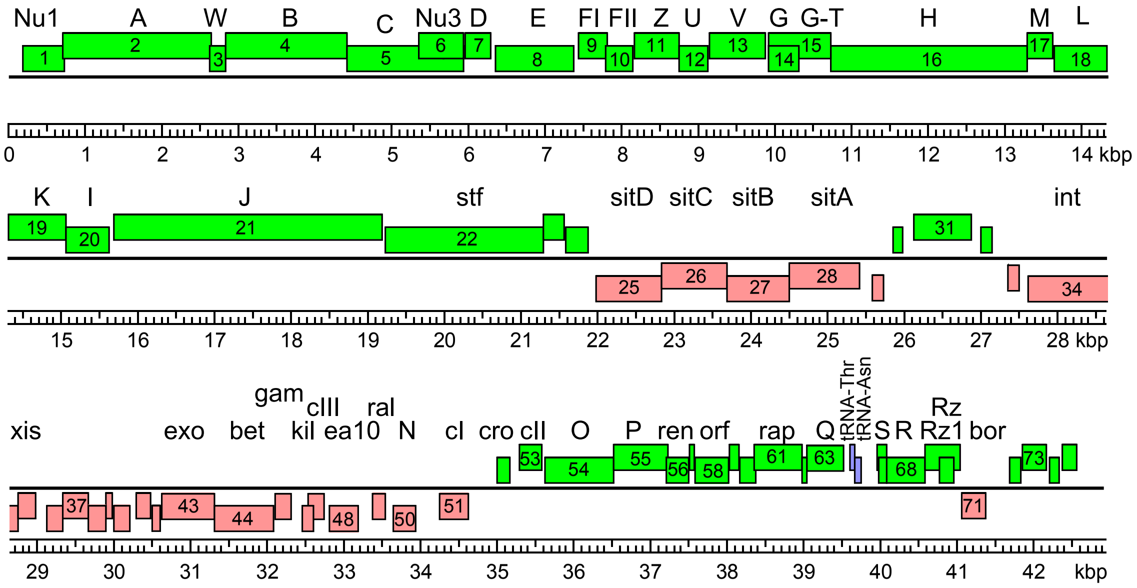
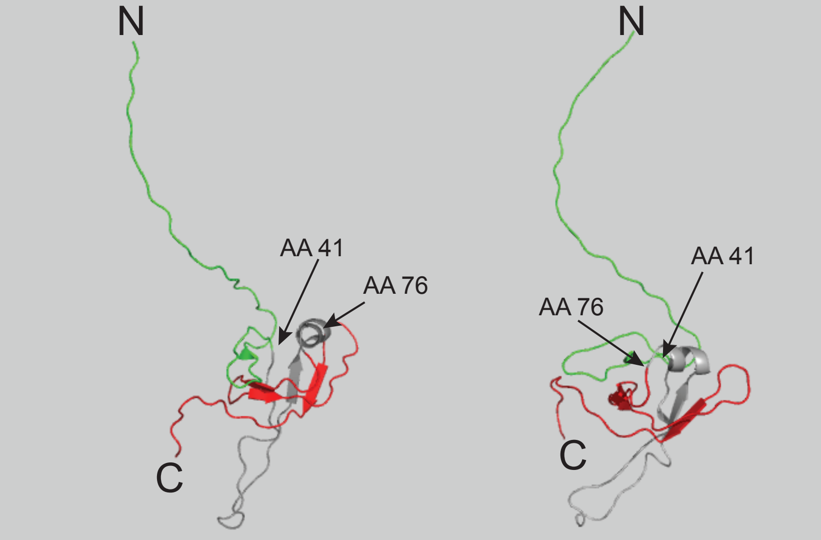


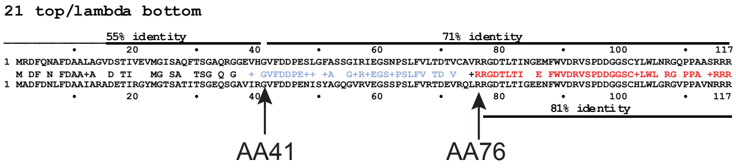
Figure S1. Map of the phage 21 genome

Colored boxes indicate predicted genes; green genes are transcribed left to right and red ones right to left; blue boxes are tRNA genes transcribed rightwards. Selected 21 gene numbers are indicated on the genes and other names including phage λ homologue names are shown above the genes. The genome is shown as a physical map of the virion chromosome, and a kbp scale is shown below the map.

Figure S2

Phage 21 gpFII mosaic boundary





```
21 top/lambda bottom

                55% identity                                 71% identity
           .        20         .        40         .        60         .        80         .       100         .    117
1 MRDFQNAFDAALAGVDSTIVEVMGISAQFTSGAQRGGEVHGVFDDPESLGFASSGIRIEGSNPSLFVLTDTVCAVRRGDTLTINGEMFWVDRVSPDDGGSCYLWLNRGQPPAASRRR
  M DF N FDAA+A  D TI   MG SA  TSG Q G   +GVFDDPE++ +A  G+R+EGS+PSLFV TD V  +RRGDTLTI  E FWVDRVSPDDGGSC+LWL RG PPA +RRR
1 MADFDNLFDAAIARADETIRGYMGTSATITSGEQSGAVIRGVFDDPENISYAGQGVRVEGSSPSLFVRTDEVRQLRRGDTLTIGEENFWVDRVSPDDGGSCHLWLGRGVPPAVNRRR
           .        20         .        40         .        60         .        80         .       100         .    117
                                                                                             81% identity
```

Figure S2. Phage 21 gpFII mosaic boundary

Two views of a ribbon diagram of the lambda gpFII protein are shown above (Maxwell *et al.*, 2002); AAs 1-40 are green; 41-76 gray; 77-C-terminus red. An alignment of the two proteins is shown below.

The differntial similarity boundary discussed in the text between low and high lambda-21 similarity is at AA 41. AA 41 lies near the boundary between the extended N-terminal region and the compactly folded domain that makes up the rest of the gpFII structure.

Figure S3

Phage 434 genome

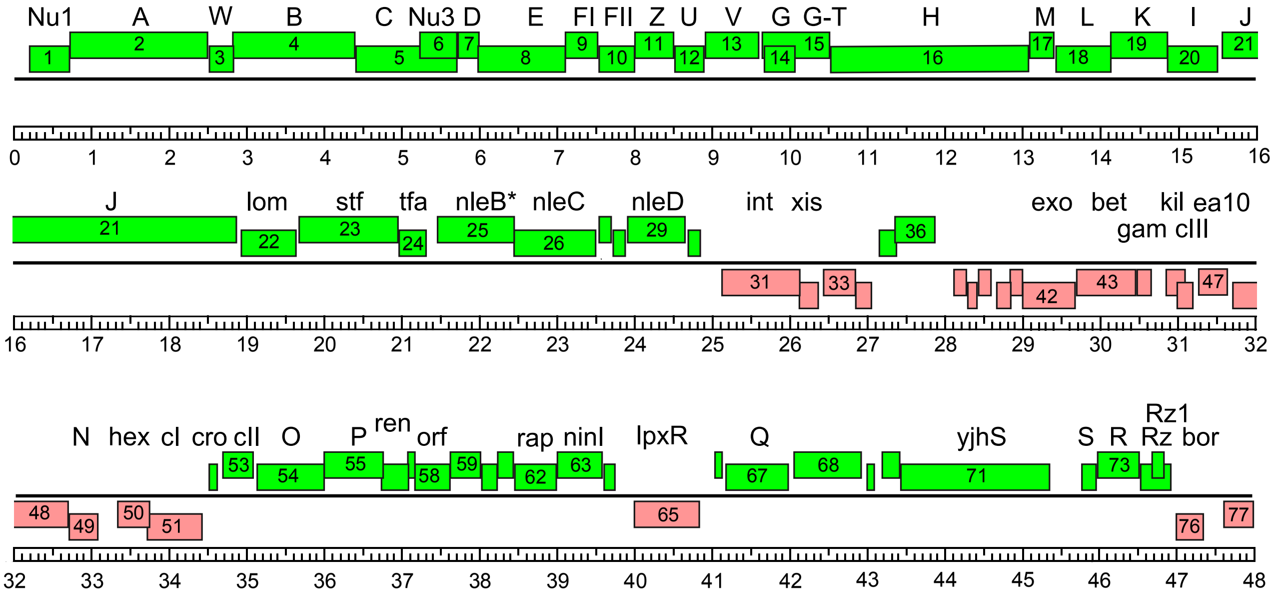
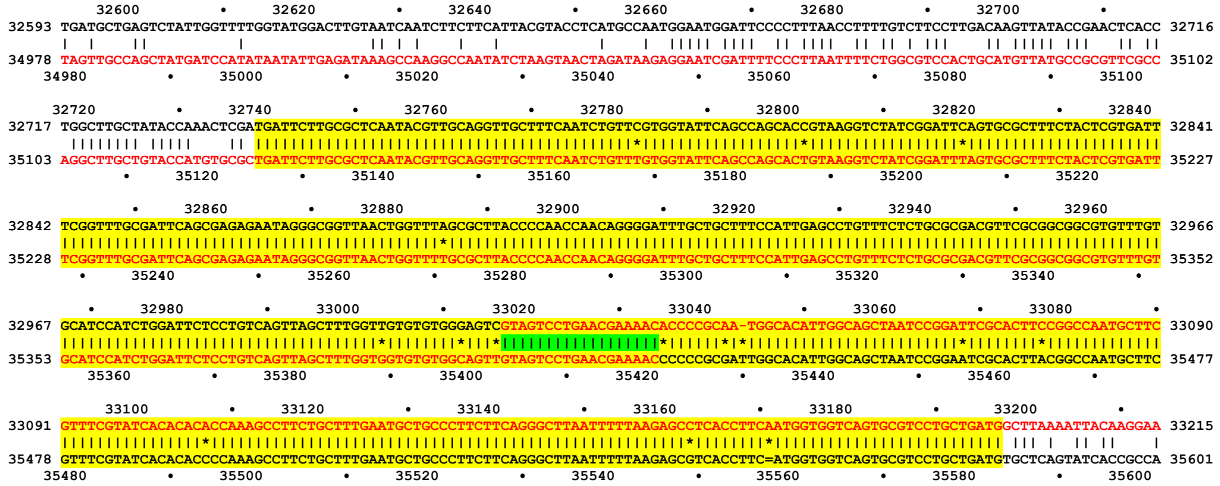


Figure S3. Map of the phage 434 genome

Colored boxes indicate predicted genes; green genes are transcribed left to right and red ones right to left. Selected 434 gene numbers are indicated on the gene and other names including phage λ homologue names are shown above the genes. The asterisk (*) indicates that the gene apparently has a broken reading frame. The genome is shown as a physical map of the virion chromosome, and a kbp scale is shown below the map.

Figure S4

434 top – lambda bottom



```
           32600       •       32620       •       32640       •       32660       •       32680       •       32700       •
32593 TGATGCTGAGTCTATTGGTTTTGGTATGGACTTGTAATCAATCTTCTTCATTACGTACCTCATGCCAATGGAATGGATTCCCCTTTAACCTTTTGTCTTCCTTGACAAGTTATACCGAACTCACC 32716
      |  |    ||          |              || | |    |   |           |    |  ||||| ||||  || ||||  || || | |||    || ||||| |||   || ||
34978 TAGTTGCCAGCTATGATCCATATAATATTGAGATAAAGCCAAGGCCAATATCTAAGTAACTAGATAAGAGGAATCGATTTTCCCTTAATTTTCTGGCGTCCACTGCATGTTATGCCGCGTTCGCC 35102
      34980       •       35000       •       35020       •       35040       •       35060       •       35080       •       35100

       32720       •       32740       •       32760       •       32780       •       32800       •       32820       •       32840
32717 TGGCTTGCTATACCAAACTCGATGATTCTTGCGCTCAATACGTTGCAGGTTGCTTTCAATCTGTTCGTGGTATTCAGCCAGCACCGTAAGGTCTATCGGATTCAGTGCGCTTTCTACTCGTGATT 32841
       |||||||| |||||    || |||||||||||||||||||||||||||||||||||||||||||*||||||||||||||||||*|||||||||||||||||*||||||||||||||||||||||
35103 AGGCTTGCTGTACCATGTGCGCTGATTCTTGCGCTCAATACGTTGCAGGTTGCTTTCAATCTGTTTGTGGTATTCAGCCAGCACTGTAAGGTCTATCGGATTTAGTGCGCTTTCTACTCGTGATT 35227
             •       35120       •       35140       •       35160       •       35180       •       35200       •       35220

              •       32860       •       32880       •       32900       •       32920       •       32940       •       32960
32842 TCGGTTTGCGATTCAGCGAGAGAATAGGGCGGTTAACTGGTTTAGCGCTTACCCCAACCAACAGGGGATTTGCTGCTTTCCATTGAGCCTGTTTCTCTGCGCGACGTTCGCGGCGGCGTGTTTGT 32966
      |||||||||||||||||||||||||||||||||||||||||||*|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
35228 TCGGTTTGCGATTCAGCGAGAGAATAGGGCGGTTAACTGGTTTTGCGCTTACCCCAACCAACAGGGGATTTGCTGCTTTCCATTGAGCCTGTTTCTCTGCGCGACGTTCGCGGCGGCGTGTTTGT 35352
        •       35240       •       35260       •       35280       •       35300       •       35320       •       35340       •

         •       32980       •       33000       •       33020       •       33040        •       33060       •       33080       •
32967 GCATCCATCTGGATTCTCCTGTCAGTTAGCTTTGGTTGTGTGTGGGAGTCGTAGTCCTGAACGAAAACACCCCGCAA-TGGCACATTGGCAGCTAATCCGGATTCGCACTTCCGGCCAATGCTTC 33090
      ||||||||||||||||||||||||||||||||||||*||||||||*|||*||||||||||||||||||*||||||*|*||||||||||||||||||||||||*||||||||*|||||||||||||
35353 GCATCCATCTGGATTCTCCTGTCAGTTAGCTTTGGTGGTGTGTGGCAGTTGTAGTCCTGAACGAAAACCCCCCGCGATTGGCACATTGGCAGCTAATCCGGAATCGCACTTACGGCCAATGCTTC 35477
           35360       •       35380       •       35400       •       35420       •       35440       •       35460       •

             33100       •       33120       •       33140       •       33160       •       33180       •       33200       •
33091 GTTTCGTATCACACACACCAAAGCCTTCTGCTTTGAATGCTGCCCTTCTTCAGGGCTTAATTTTTAAGAGCCTCACCTTCAATGGTGGTCAGTGCGTCCTGCTGATGGCTTAAAATTACAAGGAA 33215
      ||||||||||||||||*||||||||||||||||||||||||||||||||||||||||||||||||||||||*||||||||*|||||||||||||||||||||||||| ||| |  || ||    |
35478 GTTTCGTATCACACACCCCAAAGCCTTCTGCTTTGAATGCTGCCCTTCTTCAGGGCTTAATTTTTAAGAGCGTCACCTTC=ATGGTGGTCAGTGCGTCCTGCTGATGTGCTCAGTATCACCGCCA 35601
      35480       •       35500       •       35520       •       35540       •        35560       •       35580       •       35600
```

Figure S4. Location of the left crossover point in the generation of λ *imm434*

The aligned sequences of the 434 (above) and λ (below) regions immediately to the left of their nonhomologous immunity regions are shown. The 459 bp section of high similarity is highlighted in yellow, and the 15 single nucleotide differences in this region are marked by asterisks (*). The red nucleotides correspond to the λ *imm434* sequence determined in this study. The recombination event that created the left end of the 434 DNA in λ *imm434* must have happened within the region of identity highlighted in green.

Figure S5

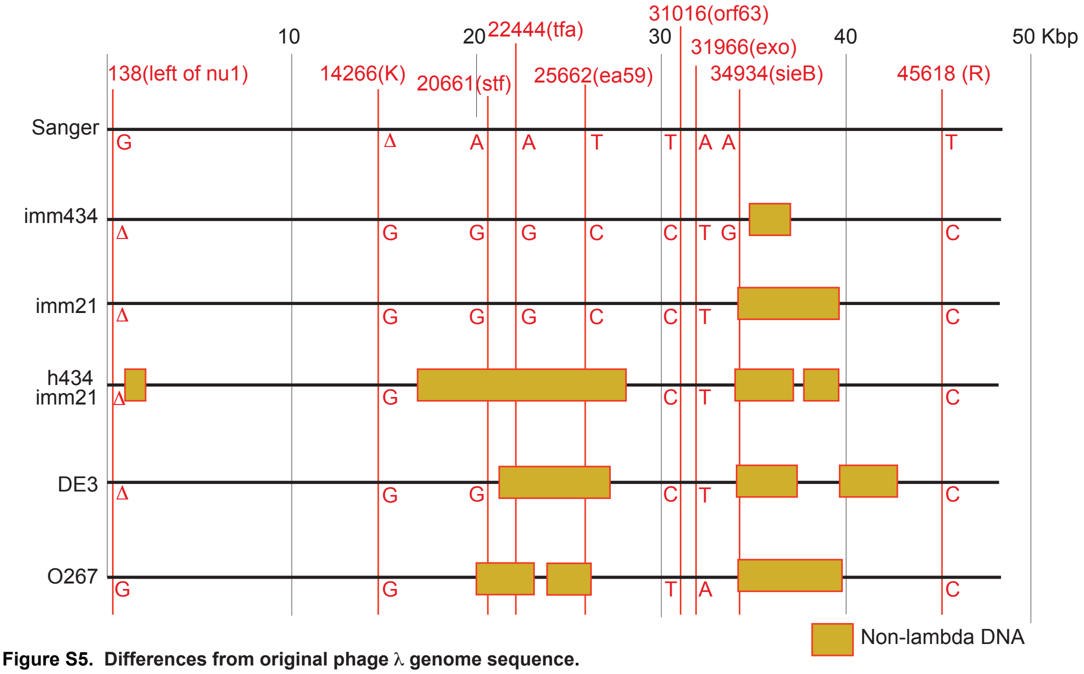

Differences from original phage λ genome sequence

Figure S5. Differences from original phage λ genome sequence.
The nine uniform differences between the three λ hybrids sequenced in this report and the original phage λ genome sequence (Acc. No. J02459) are indicated by vertical red lines and associated nucleotides in red text. In addition, the nucleotides at these locations in the sequenced λ DE3 cloning vector (Acc. No. EU078592) and HK022 hybrid O276 (Acc. No. MH547045) are shown below. Two of these nucleotides at bp 138 and 14266 were found carry the hybrid phage nucleotides in λ phages used in the laboratories of M. Feiss (unpublished) and A. Davidson (personal communication), respectively. The latter is a frameshift in the esentail gene *K*, and so may be an error in the original λ sequence; the nucleotide at 45818 in gene *R* was found to be the same as the original λ sequence in phage λ in the laboratory of R. Young (unpublihsed).

SUPPLEMENATRY MATERIAL REFERENCES

Hendrix, R.W. and R.L. Duda, *Bacteriophage lambda PaPa: not the mother of all lambda phages.* Science, 1992. **258**:1145-8.

Maxwell, K., A. Yee, C. Arrowsmith, M. Gold, and A. Davidson, *The solution structure of the bacteriophage lambda head-tail joining protein, gpFII.* Journal of Molecular Biology, 2002. **318**:1395-1404.

Sanger, F., et al., *Nucleotide sequence of bacteriophage λ DNA.* Journal of Molecular Biology, 1982. **162**:729-773.

Simon, M.N., R.W. Davis, and N. Davidson, *Heteroduplexes of DNA molecules of lambdoid phages: Physical mapping of their base sequence relationships by electron microscopy*, in *The Bacteriophage Lambda*, A.D.-. Hershey, Editor. 1971, Cold Spring Harbor Press: Cold Spring Harbor, New York. pp313-328.