J Mol Evol (1996) 42:537–542

JOURNAL OF MOLECULAR EVOLUTION



# Phylogenetic Analysis of the Isopenicillin-N-Synthetase Horizontal Gene Transfer

**Celia Buades, Andrés Moya**

Departamento de Genética, Facultad de Biología and Servicio de Bioinformática, Universidad de Valencia, c/o Dr. Moliner, 50. 46100 Burjassot, Valencia, Spain

Received: 27 May 1995 / Accepted: 2 November 1995

**Abstract.** A phylogenetic study of the isopenicillin-N-synthetase (IPNS) gene sequence from prokaryotic and lower eukaryotic producers of β-lactam antibiotics by means of a maximum-likelihood approach has been carried out. After performing an extensive search, rather than invoking a global molecular clock, the results obtained are best explained by a model with three rates of evolution. Grouped in decreasing order, these correspond to *A. nidulans* and then to the rest of the eukaryotes and prokaryotes, respectively. The estimated branching date between prokaryotic and fungal IPNS sequences (852 ± 106 MY) strongly supports the hypothesis that the IPNS gene was horizontally transferred from bacterial β-lactam producers to filamentous fungi.

**Key words:** Horizontal gene transfer — β-lactam producers — Maximum likelihood — Molecular clock

## Introduction

Horizontal gene transfers are difficult to prove, and various lines of indirect evidence are necessary to demonstrate a convincing case (Lawrence and Hartl 1992; Smith et al. 1992; Kidwell 1993; Dyer and Obar 1994). On a phylogenetic time scale, horizontal transfer will produce gene trees that do not correspond with well-established species trees, giving inconsistencies in taxonomic relationships (Lawrence and Hartl 1992; Smith et al. 1992; Kidwell 1993; Syvanen 1994).

However, the expected (i.e., species) phylogeny is not a sufficient condition to support the assumption that horizontal transfer has not occurred, specially if it took place a long time ago. Less dramatic changes than a distorted phylogeny can be the consequence of horizontal gene transfer, e.g., unexpected short branch lengths in a true species topology.

However, another problem is to ascertain whether the sequences compared are orthologous (i.e., genes derived from a common ancestor), paralogous (i.e., duplicated), or xenologous (i.e., horizontally transferred). According to Fitch (1970) and Gray and Fitch (1983) the topology of a set of taxa can be correctly obtained from orthologous genes, but paralogous or xenologous genes are not able to yield the correct species topology.

Most prokaryotic organisms are able to produce β-lactam antibiotics, whereas in eukaryotes, the ability to make these compounds is restricted to filamentous fungi (Elander 1983; Jensen 1986; Cohen 1990). In the metabolic pathway leading to such antibiotics, isopenicillin-N-synthetase (IPNS) catalyzes the transformation of the tripeptide (L-aminoadipyl)-L-cysteinyl-D-valine into isopenicillin-N. In spite of the large phylogenetic distance between eubacteria and fungi (Hori and Osawa 1987), both the pathway leading to β-lactam antibiotics and the gene sequence of IPNS are remarkably similar in these organisms. Based on these and other arguments (Smith et al. 1990a–c; Coque et al. 1991a,b), several authors have proposed the possible horizontal transfer of

*Correspondence to:* A. Moya

**Table 1.** Bacteria and fungi species, and accession numbers for the 5S rRNA and IPNS sequences

| Gene | Code | Species | Accession number |
|---|---|---|---|
| IPNS | Fla | *Flavobacterium* sp., Gram (−) | X17355 |
| | Scl | *Streptomyces clavuligerus,* Gram (+) | M19421 |
| | Sli | *Streptomyces lipmanii,* Gram (+) | M22081 |
| | Sju | *Streptomyces jumonjinensis,* Gram (+) | M36687 |
| | Ach | *Acremonium chrysogenum,* fungus | X03148 |
| | Pch | *Penicillium chrysogenum,* fungus | M15083 |
| | Ani | *Aspergillus nidulans,* fungus | M18111 |
| 5S rRNA | Afa | *Alcaligenes faecalis,* Gram (−) | X05517 |
| | Sgr | *Streptomyces griseus,* Gram (+) | J01886 |
| | Ach | *Acremonium chrysogenum,* fungus | X00867 |
| | Pch | *Penicillium chrysogenum,* fungus | X00692 |
| | Ani | *Aspergillus nidulans,* fungus | X00688 |

the IPNS genes, and probably also other genes involved in the synthesis of β-lactam antibiotics, from prokaryotic producers to filamentous fungi (Ramón et al. 1987; Miller and Ingolia 1989; Peñalva et al. 1990; Landan et al. 1990).

In the present paper we applied a statistical procedure to ascertain the case of horizontal gene transfer. We have taken advantage of the maximum-likelihood method developed by Hasegawa and co-workers (Hasegawa et al. 1985; Kishino and Hasegawa 1990) to estimate, with standard errors, evolutionary rates and split dates of a given tree. The method allows one to compare alternative models of evolution (for instance, constant vs variable rates of nucleotide substitution). These estimates are relevant to validate the horizontal gene transfer hypothesis for the IPNS gene.

## Materials and Methods

*Sequences and Alignment Algorithm.* Two different gene sequences have been used: IPNS and 5S rRNA. Table 1 shows accession numbers of IPNS and 5S rRNA sequences of the species under study. Gene sequences of 5S rRNA have been aligned following the hierarchical clustering algorithm, CLUSTAL V, of Higgins and Sharp (1989). The IPNS alignment has been obtained with the same algorithm but on the deduced amino-acid sequences and then back translated to nucleotide sequences.

*Phylogenetic Procedures and Statistical Analysis.* The programs FITCH, KITSCH, DNAML, and DNAMLK from the PHYLIP package (Felsenstein 1990), were used to test the hypothesis of a global molecular clock.

To estimate branching dates and evolutionary rates, as well as to test different models (i.e., models that assume varying evolutionary rates between branches), we followed the maximum likelihood approach (Hasegawa et al. 1985; Kishino and Hasegawa 1990) implemented in the MClock program developed by Dr. Hasegawa.

Based on the information theory, Akaike (1974) derived a criterion for the comparison of non-nested models which is very useful when comparing different tree topologies or evolutionary models with varying rates. The AIC (Akaike Information Criterion) statistic for a given model is defined as AIC = −2 (estimated log-likelihood) + 2 (number of free parameters) and it provides the basis for model selection. The better the fit of the model to the data, the lower the first term is. On the other hand, the more complex the model, the higher the second term is. The model with minimum AIC is considered to be the most appropriate model. Using this criterion, it can be determined whether or not additional parameters should be introduced in a given model.

The maximum-likelihood approach assumes independence and homogeneity among sites. However, this assumption may not be strictly valid because the composition of nucleotides may not be homogeneous along the sequence; furthermore, there may be a correlation between sites. Even if the assumption of homogeneity and independence is violated, the estimates of branching dates and the average substitution rates may still be good approximations. Nevertheless, the standard errors of the estimates should be modified, because the variance-covariance matrix is calculated under the assumption of a multinomial distribution for the frequencies of the states. We offer an approach that introduces an additional parameter that modifies the variance-covariance matrix with a correction factor. Under this modification, the estimates themselves do not change, but the standard errors are modified. The introduction of the new parameter constitutes another model, and its AIC is denoted by AIC(2) (see Kishino and Hasegawa 1990 for more details).

## Results

The aligned 5S rRNA and IPNS sequences are not shown but they can be obtained upon request from the senior author. In order to avoid the well-known saturation effect, third positions were excluded in the phylogenetic analysis of IPNS sequences. The number $k$ of nucleotide substitutions per site between pairs of sequences, following Kimura's two-parameter method (Kimura 1983), is shown in Table 2. The average values of $k$ between bacteria and fungi, within bacteria, and within fungi for IPNS and 5S rRNA are shown in Table 3. As can be observed, while there is a notorious difference between fungi and bacteria according to 5S rRNA, the distance separating both groups is much more reduced according to IPNS.

The phylogenetic trees obtained with PHYLIP programs (v3.4) based on parsimony (DNAPARS), distances (FITCH), or maximum likelihood (DNAML and DNAMLK) yielded different topologies for each gene. All methods gave us the same results: Whereas Gram (+) and Gram (−) bacteria cluster according to 5S rRNA, the Gram (+) bacteria cluster first with fungi according to IPNS (Fig. 1). It is evident that the branch connecting nodes 1 and 3 in the IPNS tree should be considered as a zero branch length and that what we have is an unresolved trifurcation among Gram (+), Gram (−), and fungi.

The evolutionary clock hypothesis was tested by means of the log-likelihood ratio test, combining results of both DNAML and DNAMLK programs. The 5S rRNA yielded a difference between log-likelihoods of 1.39, a nonsignificant value at the level of 0.05 for a $\chi^2$ with 3 degrees of freedom. The difference between the log-likelihoods for the IPNS gene in both trees was 14.40, a significant $\chi^2$ at the 0.05 level and for 5 degrees of freedom. If we consider that the 5S rRNA tree repro-

**Table 2.** Genetic distances between pairs of 5S rRNA and IPNS sequences (third positions excluded) expressed as nucleotide differences per site[a]

| Sequence | | Fla | Scl | Sli | Sju | Ach | Pch | Ani |
|---|---|---|---|---|---|---|---|---|
| IPNS | Fla | – | | | | | | |
| | Scl | 0.3574 | – | | | | | |
| | Sli | 0.4029 | 0.1914 | – | | | | |
| | Sju | 0.3533 | 0.1374 | 0.1892 | – | | | |
| | Ach | 0.3975 | 0.4085 | 0.3764 | 0.3479 | – | | |
| | Pch | 0.4144 | 0.4250 | 0.4105 | 0.3849 | 0.1709 | – | |
| | Ani | 0.4343 | 0.3936 | 0.3903 | 0.3875 | 0.1938 | 0.1488 | – |
| 5S rRNA | | Afa | Sgr | Ach | Pch | Ani | | |
| | Afa | – | | | | | | |
| | Sgr | 0.5352 | – | | | | | |
| | Ach | 1.1121 | 1.1817 | – | | | | |
| | Pch | 0.8349 | 1.0739 | 0.3213 | – | | | |
| | Ani | 0.9516 | 1.1869 | 0.3119 | 0.0911 | – | | |

[a] The distances are calculated with Kimura's two-parameter method. See Table 1 for the species names

duces the species tree (Hori and Osawa 1987), then two remarks can be made on the evolution of IPNS when compared with the evolution of 5S rRNA. First, branches connecting nodes 1 and 2 and 1 and 3 in the IPNS tree are clearly shorter than those in the 5S rRNA tree. If we use 1,000 MY as the reference and fixed time when Gram (+) and Gram (−) bacteria diverged (Hori and Osawa 1987 and node 3 of the 5S rRNA tree, Fig. 1) then, according to the IPNS tree, the splitting of node 1 should have taken place 996 MY ago, a value that is not congruent with the one reported for the species tree (2,400 MY, Hori and Osawa 1987). Second, the node 4 event is older than expected when compared with that corresponding to the species tree. According to Hori and Osawa (1987), *Penicillium* and *Aspergillus* diverged about 60 MY ago, but in the IPNS the divergence approximately occurred 282 MY ago. These two differences deserve more attention and have been tested separately by means of the maximum likelihood approach (Hasegawa et al. 1985).

The observed number of transitions and transversions between each pair of sequences is shown in Table 4. From these data, we estimated the initial values for transition ($\alpha$) and transversion ($\beta$) rates. The estimation was performed by trial and error using the expressions for $S/n$ and $V/n$ to fit the observed data (equations 12 and 13 from Kishino and Hasegawa 1990, respectively). After covering a range for $\alpha$ and $\beta$ initial estimates as wide as possible, we finally used values of 0.8 and 0.5 $MY^{-1}$, respectively (Fig. 2). As initial estimates of divergence times, we have used those obtained by transforming distance to time according to the DNAMLK tree and those obtained via a fixed divergence time between bacteria Gram (+) and Gram (−) of 1000 MY (T3, constant rate model). Other relevant parameters to the model are the total length of aligned sequences ($n$ = 682 nucleotides, gaps and third positions excluded), the proportion of variable sites (0.513), and the base composition (T = 0.220, C = 0.248, A = 0.290 and G = 0.242). The maximum-likelihood estimates with those initial values are given in the molecular clock model (Table 5). According to this model, with an AIC(2) value of 205.4, the divergence between IPNS of bacteria and fungi took place 975 ± 86 MY ago.

It should be pointed out that IPNS did not show a global clock behavior when running DNAML and DNAMLK programs. So, the second event to be tested is the high evolutionary rates of branches 7 and 8 yielding IPNS of *Penicillium* and *Aspergillus* (Fig. 2). To check this hypothesis, initial values of $\alpha$ = 0.8 and $\beta$ = 0.5 were used, except for branches 7 and 8 where initial values of $\alpha > 0.8$ and $\beta > 0.5$ were explored. As initial

**Table 3.** Average number $k$ of nucleotide substitutions per site within and between sets of species

| Species group compared | IPNS | 5s rRNA |
|---|---|---|
| Bacteria | 0.2719 | 0.5352 |
| Fungi | 0.1712 | 0.2412 |
| Bacteria vs fungi | 0.3976 | 1.0568 |

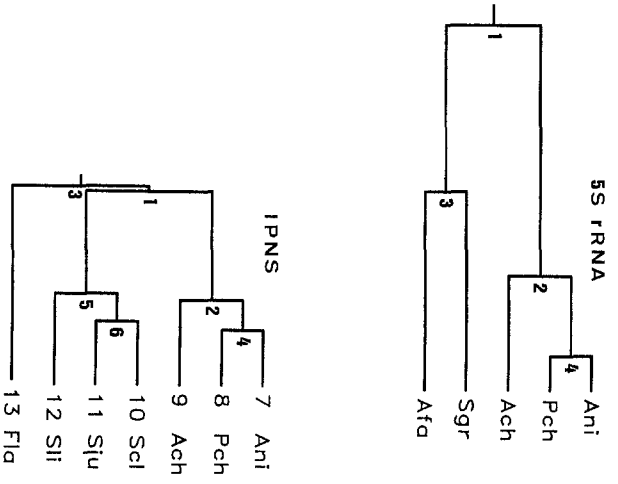


**Fig. 1.** DNAMLK trees of 5S rRNA and IPNS sequences.

**Table 4.** Number of transitions (above diagonal) and transversions (below diagonal) for pairs of IPNS sequences[a]

| | Fla | Scl | Sli | Sju | Ach | Pch | Ani |
|---|---|---|---|---|---|---|---|
| Fla | – | 60 | 66 | 62 | 67 | 64 | 70 |
| Scl | 102 | – | 42 | 27 | 70 | 81 | 68 |
| Sli | 122 | 65 | – | 42 | 67 | 72 | 62 |
| Sju | 109 | 53 | 64 | – | 70 | 82 | 73 |
| Ach | 119 | 121 | 113 | 104 | – | 41 | 45 |
| Pch | 126 | 116 | 113 | 111 | 57 | – | 38 |
| Ani | 127 | 117 | 114 | 120 | 64 | 49 | – |

[a] See Table 1 for the species names

times we used those obtained according to the molecular clock model and T4 of 60 MY (i.e., the splitting time between *Aspergillus* and *Penicillium* according to Hori and Osawa 1987). A two-rate model was additionally tested, in which we maintained the initial divergence times used for the molecular clock model, but considering an initial T4 time of 60 MY. The best result found with this hypothesis is summarized in the two-rate model (Table 5). This model [with an AIC(2) of 204.7] is better than the first one. In summary, the best ML model we have found until now is a two-rate model compatible with an IPNS gene transference that took place 945 MY ago.

A new model combining the two previous peculiarities in the IPNS evolution [i.e., horizontal gene transfer from bacteria Gram (+) to fungi and fast evolution of this gene in *Aspergillus* and *Penicillium*] with a third one has been tested. After horizontal gene transfer there should be an adaptation of the prokaryotic IPNS gene to the expression machinery of eukaryotic genomes. The effect on base composition of a prokaryotic gene transferred to a new eukaryotic environment might increase the rate of nucleotide substitutions in the branch connecting nodes 1 and 2 (leading to fungi) when compared to the other branches. Such a model incorporates three initial rates: $\alpha_2$, $\beta_2$ for the branch 1–2, $\alpha_3$ $\beta_3$ for branches 7 and 8, and $\alpha_1$, $\beta_1$ for the rest ($\alpha_1 = 0.8$, $\beta_1 = 0.5$ and $\alpha_3 \gg \alpha_2 > \alpha_1$, $\beta_3 \gg \beta_2 > \beta_1$ were explored). As initial estimates of time we chose the same values as used to obtain the molecular clock model, except that T4 was 60 MY. After a long search of initial conditions we were not able to get a three-rate model with an AIC(2) lower than the two previous models. We took advantage of Ramon et al. (1987) where they reported the very high evolutionary rate of *Aspergillus,* even higher than the rate of *Penicillium.* To avoid a four-rate model we have explored initial equal rates for branches 2 (i.e., leading to fungi; see above) and 8 (leading to *Penicillium*) with $\alpha_2$ and $\beta_2$ higher than 0.5 and 0.8, respectively, and the highest initial rate for branch 7 leading to *Aspergillus* ($\alpha_3$ and $\beta_3$ even higher than $\alpha_2$ and $\beta_2$). The initial times were the same as used for two-rates model. As can be observed, the AIC(2) is 201.0, the lowest obtained in our search.

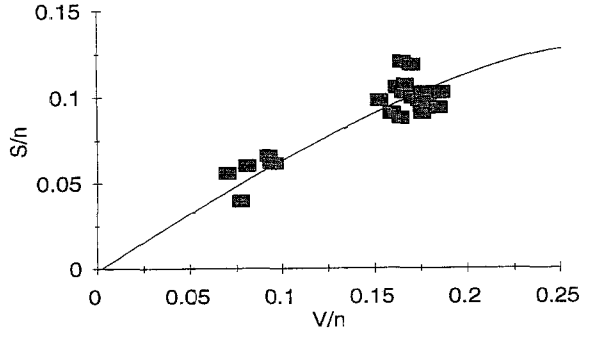


**Fig. 2.** Relationship between the number of transitions (*S/n*) and transversions (*V/n*) per nucleotide site for the IPNS sequences used (gaps and third positions excluded). The α and β estimates that better fit *S* and *V* to empirical values of transitions and transversions (Table 4) are 0.8 and 0.5, respectively.

## Discussion

The IPNS tree obtained cannot solve the trifurcation among Gram (+), Gram (−), and fungi. This situation, itself, indicates that this gene does not reproduce the

**Table 5.** Branching dates (MY) and evolutionary rates estimated from IPNS sequences

| | Model | | |
|---|---|---|---|
| Parameter | Molecular clock | Two rates | Three rates |
| Branching dates | | | |
| T1 | 975 ± 86 | 945 ± 90 | 852 ± 106 |
| T2 | 356 ± 43 | 249 ± 68 | 316 ± 40 |
| T3 | 1000 | 1000 | 1000 |
| T4 | 266 ± 37 | 47 ± 121 | 146 ± 41 |
| T5 | 395 ± 46 | 408 ± 48 | 423 ± 45 |
| T6 | 286 ± 40 | 297 ± 41 | 310 ± 39 |
| Rates | | | |
| $\alpha_1$ | 0.936 ± 0.136 | 0.859 ± 0.125 | 0.797 ± 0.113 |
| $\beta_1$ | 0.661 ± 0.081 | 0.645 ± 0.079 | 0.628 ± 0.074 |
| $\alpha_2$ | – | 6.295 ± 16.151 | 1.639 ± 0.559 |
| $\beta_2$ | – | 3.532 ± 9.084 | 0.918 ± 0.248 |
| $\alpha_3$ | – | – | 2.413 ± 0.946 |
| $\beta_3$ | – | – | 1.419 ± 0.532 |
| AIC(2) | 205.4 | 204.7 | 201.0 |

species tree (see Fig. 1). However, in our opinion, we are not dealing with a topological question but with a branch length problem. It is very difficult to explain, if one does not return to the horizontal transfer hypothesis, why the eukaryotic IPNS sequences appear so close to the prokaryotic ones, approximately 1400 MY closer than expected according to the species tree.

The maximum-likelihood analysis carried out shows that prokaryotic and eukaryotic IPNS genes did not diverge with the species but, on the contrary, were first present in bacterial β-lactam producers and transferred from bacteria to the ancestor of filamentous fungi about 852 MY ago. Another finding, in addition to the transference, is the faster evolution observed in the branches leading to *Aspergillus* and *Penicillium,* if compared to the other.

Recently Smith et al. (1992) considered that the simplest interpretation concerning the evolution of IPNS is that this gene and the gene encoding deacetoxycephalosporin C synthetase (DAOCS) underwent a duplication before the prokaryote/eukaryote divergence. They claimed that the topology of the tree rooted with the duplicated enzymes, the depth of the bacterial branches, and the different orientations of the β-lactam metabolic pathway genes in fungi and eubacteria are consistent with an ordinary evolution for IPNS gene. This is contrary to our hypothesis (Peñalva et al. 1990) and the statistical analysis of phylogenetic results here reported. Being more parsimonious, obtaining a phylogeny with a similar topology to the expected does not systematically invalidate the occurrence of a horizontal transfer event. The phylogenetic reconstruction we have obtained shows such an impressive shortening of branches linking fungi and bacteria that it is difficult to explain by slow evolution of IPNS. There is no evidence of this fact when we observe the estimated evolutionary rates within either prokaryotes or fungi (Table 3).

There is a set of arguments that makes the consideration of IPNS, and probably also DAOCS, as orthologous genes very unlikely. In their Fig. 4, Smith et al. (1992) get a 57% average percent identity of amino acid sequences for both IPNS and DAOCS from fungi and eubacteria. If we translate these values to the number of nucleotide substitutions per site (Kimura's two-parameter method, Kimura 1983), the average genetic divergence per site between IPNS or DAOCS from fungi and eubacteria is 0.40. The average genetic divergence per site between fungi and eubacteria using 5S rRNA is approximately 1.10. Accordingly, the divergence between these two groups took place 2,400 MY ago. If we also assume that IPNS behaves as a molecular clock, the divergence between IPNS of fungi and bacteria took place 975 MY ago, approximately 1,425 MY after the divergence between prokaryotes and eukaryotes.

More arguments can be invoked favoring the IPNS horizontal gene transference (see Peñalva et al. 1990), but none of them as strong as the statistical one used here. The duplication hypothesis of Smith et al. (1992) does not invalidate the transference. Though 22% of amino acid identity is a very low percentage, there is evidence (Roach et al. 1995) indicating that IPNS, DAOCS, and related 2-oxo-acid-dependent oxygenases constitute a structural family of enzymes with identified homologous regions, conserved structural and active-site motifs, and, probably, reaction mechanisms closely related. It is very likely that an ancestral duplication event occurred but it also seems clear that, afterward, a transference of IPNS (and probably some other enzymes of the β-lactamic antibiotic pathway) from prokaryotes to eukaryotes occurred.

*Acknowledgments.* The work has been supported by grant PB93-0690 from DGICYT (Spain) to A.M. and by a fellowship from the University of Valencia (Spain) to C.B. The MClock program was kindly supplied by Dr. M. Hasegawa. We also thank E. Barrio, F. González-Candelas, A. Latorre, and R. van Ham for reading and helpful criticism.

## References

Akaike H (1974) A new look at the statistical model identification. Inst Electrical Electronics IEEE Trans Autom Contr AC-19:716–723

Cohen G, Shiffman D, Mavarech M, Aharonowitz Y (1990) Microbial isopenicillin N synthase genes: structure, function, diversity and evolution. Trends Biotechnol 8:105–111

Coque JJR, Martín JF, Calzada JG, Liras P (1991a) The cephamycin biosynthetic gene *pcb*AB, encoding a large multidomain peptide synthetase, and *pcb*C of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *A. chrysogenum* and *P. chrysogenum.* Mol Microbiol 5:1125–1133

Coque JJR, Liras P, Laiz L, Martín JF (1991b) A gene encoding lisyne 6-aminotransferase, which forms the β-lactam precursor α-aminoadipic acid, is located in the cluster of cephamycin biosynthetic genes in *Nocardia lactamdurans.* J Bacteriol 173:6258–6264

Dyer BD, Obar RA (1994) Tracing the history of eukaryotic cells. Columbia University Press, New York

Elander RP (1983) Strain improvement and preservation of β-lactam producing microorganisms. In: Demain AL, Solomon NA (eds) Antibiotic containing the β-lactam structure I. Springer, Berlin, p 97

Felsenstein J (1990) PHYLIP manual. Version 3.3. University of California Press, Berkeley

Fitch WM (1970) Distinguishing homologous from analogous proteins. Syst Zool 19:99–113

Gray GS, Fitch WM (1983) Evolution of antibiotic resistance genes: the DNA sequence of a kanamycin resistance gene from *Staphylococcus aureus.* Mol Biol Evol 1:57–66

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Higgins DG, Sharp PM (1989) Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl Biol Sci 5:151–153

Hori H, Osawa S (1987) Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. Mol Biol Evol 4:445–472

Jensen SE (1986) Biosynthesis of cephalosporins. Crit Rev Biotechnol 3:277–301

Kidwell M (1993) Lateral transfer in natural populations of eukaryotes. Annu Rev Genet 27:235–256

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge

Kishino H, Hasegawa M (1990) Converting distances to time: an application to human evolution. Methods Enzymol 183:550–570

Landan G, Cohen G, Aharonowitz Y, Shuali Y, Graur D, Shiffman D (1990) Evolution of isopenicillin N synthase may have involved horizontal gene transfer. Mol Biol Evol 7:399–406

Lawrence JG, Hartl DL (1992) Inference of horizontal genetic transfer from molecular data: an approach using bootstrap. Genetics 131: 753–760

Miller JR, Ingolia TD (1989) Cloning and characterization of beta-lactam biosynthetic genes. Mol Microbiol 3:689–695

Peñalva MA, Moya A, Dopazo J, Ramón D (1990) Sequences of isopenicillin N synthetase genes suggest horizontal gene transfer from prokaryotes to eukaryotes. Proc R Soc Lond [Biol] 241:161–169

Ramón D, Carramolino L, Patiño C, Sánchez F, Peñalva MA (1987) Cloning and characterization of the isopenicillin N synthetase gene mediating the formation of the β-lactam ring in *Aspergillus nidulans*. Gene 57:171–181

Roach PL, Clifton IJ, Fülöp V, Harlos K, Barton GJ, Hajdu J, Anderson I, Schofield CJ, Baldwin JE. (1995) Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature 375:700–704

Smith DJ, Burnham MKR, Bull JH, Hodgson JE, Ward JM, Browne P, Brown L, Barton B, Earl AJ, Turner G (1990a) β-lactam antibiotic biosynthetic genes have been conserved in clusters in prokaryotes and eukaryotes. EMBO J 9:741–747

Smith DJ, Burnham MKR, Edwards J, Earl AJ, Turner G (1990b) Cloning and heterologous expression of the penicillin biosynthetic gene cluster from *P. chrysogenum*. BioTechnology 8:39–41

Smith DJ, Earl AJ, Turner G (1990c) The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *P. chrysogenum* is a 421.073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. EMBO J 9:2743–2750

Smith MW, Feng DF, Doolittle RF (1992) Evolution by adquisition: the case for horizontal gene transfers. Trends Biochem Sci 17:489–493

Syvanen M (1994) Horizontal gene transfer: evidence and possible consequences. Annu Rev Genet 28:237–261