12 February 1999

Science

Vol. 283 No. 5404
Pages 893–1072 $8

American Association for the Advancement of Science

The results presented here provide a mechanistic basis for how high-affinity ethylene binding to the sensor domain of ETR1 is achieved. We propose that ethylene interacts with a Cu(I) cofactor in an electron-rich hydrophobic pocket formed by membrane-spanning helices of the ETR1 dimer. The binding site must confer some unusual chemistry on the copper ion, because the stability of this ethylene/receptor complex (half-life for dissociation = 11 hours) (*7*) is much different from that observed for artificial copper complexes (*15*). The ability of Ag(I) ions to interact with the receptor and bind ethylene is of interest because silver ions can inhibit ethylene responses when applied to plant tissues (*4*). Silver ions may occupy the binding site and interact with ethylene but fail to induce the changes in the receptor that are needed to elicit downstream signaling.

The discovery of a functional ethylene-binding domain in the slr1212 coding sequence from *Synechocystis* raises interesting questions about the evolutionary origin of the higher plant receptors. *Synechocystis* is thought to share a common ancestor with the cyanobacterial lineage that evolved into the modern chloroplast of higher plants (*16*). The presence of both the ethylene-sensor domain and histidine-kinase transmitter domains in the cyanobacterial genome may have provided the raw materials for the evolution of the higher plant form of the ethylene receptors. Sequence homology to the ethylene-binding domain has not been identified in any other bacterial genomes sequenced to date, which supports a single origin for this functional domain in the evolution of photosynthetic organisms.

References and Notes

1. L. J. Ignarro, *Biochem. Pharmacol.* **41**, 485 (1991); M. A. Gilles-Gonzalez, G. S. Ditta, D. R. Helsinki, *Nature* **350**, 170 (1991).
2. E. C. Sisler, *Plant Physiol.* **64**, 538 (1979); T. Bengochea *et al.*, *Planta* **148**, 397 (1980).
3. S. P. Burg and E. A. Burg, *Science* **148**, 1190 (1965).
4. E. M. Beyer Jr., *Plant Physiol.* **58**, 268 (1976).
5. C. Chang, S. F. Kwok, A. B. Bleecker, E. M. Meyerowitz, *Science* **262**, 539 (1993).
6. G. E. Schaller, A. N. Ladd, M. B. Lanahan, J. M. Spanbauer, A. B. Bleecker, *J. Biol. Chem.* **270**, 12526 (1995).
7. G. E. Schaller and A. B. Bleecker, *Science* **270**, 1809 (1995).
8. The ETR1(1-128)GST fusion protein consists of the first 128 amino acids of ETR1, fused to the coding region of GST obtained from the bacterial expression vector pGEX4T1 (Pharmacia). The fusion protein was expressed in yeast with the use of the expression vector pYcDE-2 (*6*).
9. F. I. Rodriguez and A. B. Bleecker, data not shown.
10. Cultures of *Saccharomyces cerevisiae* strain LRB 520 (*6*) were grown to mid-log phase at 30°C and harvested by centrifugation at 1500*g* for 5 min. Cell pellets were washed and resuspended in 50 mM tris-HCl (pH 7.4), 10% glycerol, and 1.0 mM phenylmethylsulfonyl fluoride (PMSF) at a concentration of 0.5 g of cells per milliliter. Cell suspensions were mixed with an equal volume of glass beads and disrupted in a bead beater with a cooling bath (BioSpec, Bartlesville, OK). Cell debris was separated from the homogenate by centrifugation at 10,000*g* and total cell membranes were harvested by centrifugation at 100,000*g* for 30 min and then resuspended in assay buffer [10 mM MES (pH = 5.5), 20% sucrose, 1% dimethyl sulfoxide, and 1 mM PMSF]. Membrane preparations obtained from 2 g of cells were diluted with assay buffer to 1 ml and tested for ethylene binding (*11*).
11. In vitro ethylene-binding assays were performed as described (*7*) with some modifications as follows: 500-μl aliquots of membrane suspensions representing 1 g of yeast cells were pipetted onto strips of Whatman paper 1M (2.5 cm by 20 cm) that were rolled and inserted in Eppendorf tubes. The filters were incubated for 4 hours in sealed glass chambers containing $^{14}C_2H_4$ (0.1 μl/liter) or $^{14}C_2H_4$ (0.1 μl/liter) plus $^{12}C_2H_4$ (100 μl/liter). The filters were aired for 10 min and transferred to individual sealed chambers containing 300 μl of 250 mM mercuric perchlorate in a scintillation vial; $^{14}C_2H_4$ trapped in the mercuric perchlorate was then determined as described (*7*).
12. J. Hua, C. Chang, Q. Sun, E. M. Meyerowitz, *Science* **269**, 1712 (1995); J. Q. Wilkinson, M. B. Lanahan, H. C. Yen, J. J. Giovannoni, H. J. Klee, *ibid.* **270**, 1807 (1995).
13. A. E. Hall, G. E. Schaller, H. J. Klee, A. B. Bleecker, unpublished material.
14. To construct the slr1212 knockout cassette, a genomic fragment [831 base pairs (bp)] extending from 44 bp upstream of the start codon (Kazusa DNA Research; 897309–898142 bp; cosmid cs0328) was cloned in pBC-KS (Stratagene). A filled-in Sal I fragment containing the kanamycin gene from Tn*903* was inserted into the Sca I site within the slr1212 fragment. Homologous recombination was performed as described (*17*). Genomic Southern (DNA) blots and polymerase chain reaction analysis were used to verify gene disruption in 12 independent lines. Ethylene-binding assays in *Synechocystis* cells were performed as described (*7*), except that the cells were harvested by centrifugation, resuspended in growth media, and pipetted onto the glass fiber filters. The samples were aired for 15 min after the $^{14}C_2H_4$ incubation.
15. J. S. Thompson, R. L. Harlow, J. F. Whitney, *J. Am. Chem. Soc.* **105**, 3522 (1983); M. Munakata, S. Kitagawa, S. Kosome, A. Asahara, *Inorg. Chem.* **25**, 2622 (1986).
16. A. Wilmotte, in *The Molecular Biology of Cyanobacteria*, D. A. Bryant, Ed. (Kluwer Academic, Amsterdam, 1994), pp. 1–25.
17. J. Yu, L. B. Smart, Y. S Jung, J. Golbeck, L. McIntosh, *Plant Mol. Biol.* **29**, 331 (1995); A. Wilde, Y. Churin, H. Schubert, T. Borner, *FEBS Lett.* **406**, 89 (1997).
18. D. B. Smith and K. S. Johnson, *Gene* **67**, 31 (1988).
19. GAF is an acronym for a domain present in guanosine 3′,5′-cyclic monophosphate phosphodiesterases, *Anabaena* adenylate cyclases, and *Escherichia coli* FhlA [as described by L. Aravin and C. P. Ponting, *Trends Biochem. Sci.* **22**, 458 (1997)].
20. Membrane preparations (*10*) were made with 600 μM $CuSO_4$ in the extraction buffer. Ethylene-binding activity was solubilized as described by R. Serrano [*FEBS Lett.* **156**, 11 (1983)], using a solubilization buffer consisting of 50 mM tris (pH 8.0) and 0.5% (w/v) α-lysophosphatidylcholine (LPC). Solubilized samples were purified by affinity chromatography as described (*18*) using buffers including 0.5% LPC (w/v) previously purified in Chelex-100 columns. The purified extracts were desalted through Sephadex G-25 columns.
21. Samples were treated with 100 mM dithiothreitol and separated by SDS–polyacrylamide gel electrophoresis, and immunoblots were performed as described (*6*) using antibodies to GST (Sigma).
22. PAS is an acronym for a domain present in *Drosophila* Per, mammalian Arnt, and *Drosophila* Sim [as described by I. B. Zhulin, B. L. Taylor, R. Dixon, *Trends Biochem. Sci.* **22**, 331 (1997)].
23. D. T. Jones, W. R. Taylor, J. M. Thornton, *Biochemistry* **33**, 3038 (1994).
24. Amino acid sequences were aligned by means of the clustal method with the PAM 250 residue weight table (Megalign-DNASTAR, DNASTAR, Madison, WI, 1993).

25. We thank J. Spanbauer and A. Hahr for their contributions during the initial stages of this work; J. Burstyn and members of her laboratory for their helpful advice and assistance; and C. Chang and S. E. Patterson for their comments and help on the manuscript. Supported by NSF (grant 9513463), the U.S. Department of Energy (DOE) (grant DEFG02-91ER20029), and the DOE-NSF-USDA Collaborative Research in Plant Biology Program (grant DBI960-2222).

21 October 1998; accepted 12 January 1999

A Molecular Phylogeny of Reptiles

S. Blair Hedges* and Laura L. Poling

The classical phylogeny of living reptiles pairs crocodilians with birds, tuataras with squamates, and places turtles at the base of the tree. New evidence from two nuclear genes, and analyses of mitochondrial DNA and 22 additional nuclear genes, join crocodilians with turtles and place squamates at the base of the tree. Morphological and paleontological evidence for this molecular phylogeny is unclear. Molecular time estimates support a Triassic origin for the major groups of living reptiles.

The number of temporal openings in the skull has long been viewed as a key character in the classification of reptiles (*1*, *2*). A single opening (synapsid condition) is found in mammals and their reptilian ancestors. Turtles and several late Paleozoic and early Mesozoic groups lack a temporal opening (anapsid condition). Most other living and fossil reptiles belong to a clade in which the ancestral condition was the presence of two temporal openings (diapsid condition). Among the living diapsid reptiles, the crocodilians and birds form one group, the Archosauria, and the tuataras and squamates (lizards, snakes, and amphisbaenians) form another group, the Lepidosauria. Other morphologi-

Department of Biology, Institute of Molecular Evolutionary Genetics, and Astrobiology Research Center, 208 Mueller Laboratory, Pennsylvania State University, University Park, PA 16802, USA.

*To whom correspondence should be addressed at 208 Mueller Laboratory, Pennsylvania State University, University Park, PA 16802, USA. E-mail: sbh1@psu.edu

cal characters (*2*–*7*) support this traditional phylogeny of living amniote vertebrates (Fig. 1A). An alternative phylogeny is that turtles are diapsid reptiles and most closely related to the lepidosaurs (*8*, *9*); however, this lepidosaur connection has been contested (*10*, *11*).

Molecular phylogenies have been unable to corroborate the traditional phylogeny of living amniotes. In most studies turtles have clustered with archosaurs (*12*–*14*), but the number of genes or sites available generally has been small. When only crocodilians, birds, and mammals were considered, combined sequence data from 15 nuclear genes significantly supported a bird-crocodilian relationship (*15*), but the position of turtles and other reptiles was left unresolved. Here we address relationships among the major groups of living reptiles (turtles, tuataras, squamates, crocodilians, and birds) with new evidence from two nuclear genes and analyses of other available molecular sequence data bearing on this question.

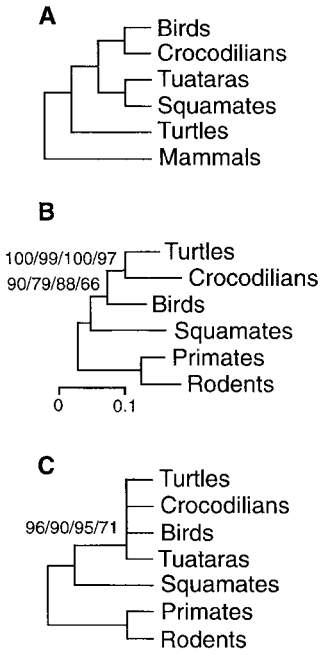

Fig. 1. Relationships among the major groups of living reptiles. (**A**) The classical phylogeny based on morphology and the fossil record (*1*, *2*). (**B**) Maximum likelihood phylogeny of combined sequences from 11 nuclear proteins (1943 amino acids). Scale bar indicates amino acid substitutions per site. (**C**) Consensus phylogeny of combined sequences from four nuclear protein-coding genes for which sequences of tuatara are available (785 amino acids). For the molecular trees, confidence values (%) supporting the nodes are separated by slash marks and based on the following four methods: interior-branch test, and bootstrap analyses of neighbor-joining, maximum likelihood, and maximum parsimony, respectively. In (B) it was determined that myoglobin had the greatest effect in lowering confidence values; removal of that gene did not change significance of turtle-crocodilian node but raised support for bird-turtle-crocodilian node to 94/90/94/88.

We sequenced the coding region of the gene for alpha enolase (phosphopyruvate hydratase) in five species of reptiles and the gene for 18*S* ribosomal RNA (rRNA) in a tuatara to compare with other reptilian sequences of those genes (*16*). For phylogenetic and timing analyses, these new sequences were added to 340 available protein and DNA sequences representing 24 nuclear and 9 mitochondrial genes (*17*). The sequences were aligned and subjected to several phylogenetic analyses (*18*).

Turtles clustered with one or both archosaurs in individual analyses of all 15 genes having representatives of squamates, birds, crocodilians, and turtles (Table 1). In nine of those genes, turtles joined crocodilians in either nucleotide or amino acid analyses. The classical phylogeny (Fig. 1A) was not supported in any comparison, and only one gene analysis resulted in a turtle-squamate relationship. A combined analysis of sequences from all nuclear proteins resulted in significant support (>97% confidence) for a turtle-crocodilian relationship on the basis of an interior branch test and bootstrapping of neighbor-joining, maximum likelihood, and maximum parsimony trees (Fig. 1B). The maximum likelihood tree (Fig. 1B) ($\ln L = -11405$) was significantly better ($\Delta \ln L > 2$ SE) than the classical tree (Fig. 1A) ($\ln L = -11481 \pm 22.1$) or the tree grouping turtles with lepidosaurs (squamates + tuataras; $\ln L = -11477 \pm 22.6$). A four-cluster analysis (*19*) also supported (>98% better than both alternative hypotheses) a turtle-crocodilian relation when combined protein sequence data (1943 sites) and combined nuclear rRNA sequence data (2299 sites) were analyzed separately.

Fewer genes are available to address the position of the tuataras, but in six of eight comparisons they cluster with archosaurs or turtles rather than squamates (Table 2). Tuataras are excluded from a close relation with squamates in all analyses of combined proteins, and this is significant with the interior branch test (96%) and bootstrap analysis of maximum likelihood (95%) (Fig. 1C). However, the difference between the maximum likelihood tree ($\ln L = -5516$) and the classical phylogeny (Fig. 1C) ($\ln L = -5540 \pm 16.5$) is not significant (<2 SE). Although the remaining nodes in the combined protein analysis are not well resolved, the combined nucleic acid analysis (Table 2) defines a cluster (turtles, crocodilians, and birds) that excludes tuataras. If 18*S* rRNA, which favors a

Table 1. Molecular evidence for the closest living relative of turtles.

Gene	Sites	Taxa	Closest relative		Basal lineage	
			Taxon* and BP%†		Taxon* and BP%‡	
Nuclear, amino acid						
A-cryst	149	15	C+B	69.0	S	69.0
A-eno	373	8	C+B	45.2	S	45.2
A-glob	141	18	B	36.5	S	94.5
B-glob	146	21	C	58.2	S	47.9
Calcit	32	7	C+B or B§	85.9	S	87.0
C-mos	125	39	C	36.4	S	32.0
Cytc	106	8	C	66.1	S	98.1
Insulin	51	9	B‖	68.5	B	56.7
Ldha	333	7	C	89.7	B	43.0
Ldhb	334	6	C	50.3	S	65.4
Myoglobin	153	10	C	86.4	B	90.8
Nuclear, nucleic acid						
18S rRNA	1947	8	C	95.2	B	98.4
28S rRNA	352	7	C	46.2	S	76.1
A-eno	1120	8	C	98.3	S	84.2
C-mos	376	40	C+B	37.4	S	37.4
Mitochondrial, amino acid						
Cytb	296	15	C+B	32.1	S	32.1
Mitochondrial, nucleic acid						
Mt-npc	3264	6	B	54.5	S	98.8
Cytb	873	18	S	43.3	S+T	42.9
Combined analyses						
Nuclear, amino acid	1943	6	C	98.9	S	79.0
Nuclear, nucleic acid (not rRNA)	1496	6	C	86.8	S	91.0
Nuclear rRNA genes	2299	6	C	92.7	B	83.9
All nuclear, nucleic acid	3795	6	C	94.4	S	85.5

*C, crocodilians; B, birds; S, squamates; T, turtles. †BP, bootstrap *P*-value in neighbor-joining analysis. ‡BP support for node joining all taxa exclusive of basal lineage. §Sequence of turtle identical to that of bird and crocodilian. ‖Sequence of turtle identical to that of some birds (turkey and ostrich), whereas goose is basal.

bird-mammal grouping (*13*), is removed, support for that cluster is significant (100%, interior branch; 95% neighbor-joining, 97% maximum likelihood). Thus the classic grouping of squamates with tuataras is not supported by available molecular data, and there is support in most analyses for tuataras as closest relatives of a group containing turtles, crocodilians, and birds.

We analyzed a total of 23 nuclear genes and the two mitochondrial regions (9 genes) to estimate divergence times (*20*). The time estimates indicate that squamates diverged from the other reptiles at 245 ± 12.2 million years ago (Ma) (9 genes), birds diverged from the lineage leading to turtles and crocodilians at 228 ± 10.3 Ma (17 genes), and that turtles diverged from crocodilians at 207 ± 20.5 Ma (7 genes). These divergence times are close to when the first turtles (223 to 210 Ma) and crocodilians (210 to 208 Ma) appear in the fossil record (*21*) and earlier than the first birds (152 to 146 Ma) and first squamates (157 to 155 Ma). These results support the notion that many key innovations in tetrapod evolution occurred during the Triassic (251 to 208 Ma) (*2*).

In light of this phylogeny of reptiles, early molecular analyses that clustered birds with mammals (*13*, *22*) now are more easily explained. When there are no lepidosaurs in an analysis, birds become the basal lineage of reptiles. Thus, birds are closer to mammals in a network and may join together more easily in some analyses, especially if rates of evolution vary among sites or lineages. Typical proteins comprise only 200 to 300 amino acids, and therefore strong statistical support for any phylogeny usually requires multiple genes.

The use of different outgroups also can result in different topologies. For example, a previous study of mitochondrial rRNA genes, in which a more distant (amphibian) root was used, yielded a bird-crocodilian grouping that excluded turtles (*15*). In this study, mammals were used to root the trees because they were the closest available outgroup and yielded the best alignments. However, the use of more distant roots did not affect the topology of most nuclear genes and did not change the conclusions of this study.

The molecular support for a turtle-crocodilian clade is surprising considering that it seems to have virtually no support from morphology. Even recent studies showing diapsid affinities of turtles did not find a close relation between turtles and archosaurs (*8*, *9*). It has been reported that turtles are most similar to crocodilians in sperm morphology (*23*), but phylogenetic analysis of sperm characters did not support that proposition (*24*).

Turtles and crocodilians are the only living tetrapod groups with dorsal and ventral bony plates of armor. Some lizards have superficial dermal ossifications, but there is no evidence of dermal armor in primitive squamates, sphenodontians, or synapsids (*25*). Ventral armor is known in extinct diapsids, including sauropterygians and archosauromorphs (*25*), but the significance of this character will be unclear until the position of turtles and crocodilians among those groups is established.

Crocodilians have large heads, long snouts, and well-developed teeth. However, some Triassic suchians (archosaurs), such as the aetosaurs (*2*), have small heads with beaklike jaws and greatly reduced teeth. Body armor was well developed, and their ventral plating has been described as a plastron (*2*, *25*). In one aetosaur (*25*), the neck spines resembled those of an early turtle (*26*). Some or all of these similarities may be the result of convergence, but they show that characteristics of turtles and crocodilians can be found together in some extinct reptiles of the Triassic.

The finding in most analyses of the combined molecular data (Table 2 and Fig. 1C) that tuataras and squamates are not closely related also is unconventional. Recently, an analysis of characters from sperm morphology argued for a dismantling of the Lepidosauria (*24*), but otherwise the phylogeny in that study does not agree with the results of this study.

These results highlight a significant discordance between morphological and molecular estimates of phylogeny for a major group of organisms. The consistent pattern of a turtle-crocodilian relationship across independent nuclear genes stands in contrast to the traditional phylogeny of amniote vertebrates. Determining how the many groups of extinct reptiles of the late Paleozoic and early Mesozoic fit into this molecular phylogeny will be a challenge to paleontologists.

Table 2. Molecular evidence for the closest living relative of tuataras.

Gene	Sites	Taxa	Closest relative		Basal lineage	
			Taxon and BP%		Taxon and BP%	
Nuclear, amino acid						
A-eno	373	9	S	40.1	S+T	35.9
A-glob	141	19	B	80.0	S	93.8
B-glob	146	22	B	24.7	S	47.1
C-mos	125	40	T	62.4	S	24.0
Nuclear, nucleic acid						
18S rRNA	1873	9	T	47.4	B	97.7
A-eno	1120	9	T+B+C	61.1	S	61.1
C-mos	376	41	S	60.2	B+C	64.8
Mitochondrial, nucleic acid						
Mt-npc	2861	8	T+B+C	89.7	S	89.7
Combined analyses						
Nuclear, amino acid	785	7	B	35.3	S	89.8
All nucleic acid	5243	7	T+B+C	76.2	S	76.2

References and Notes

1. A. S. Romer, *Vertebrate Paleontology* (Univ. of Chicago Press, Chicago, IL, 1966).
2. M. J. Benton, *Vertebrate Paleontology* (Chapman & Hall, New York, ed. 2, 1997).
3. M. Laurin and R. R. Reiz, *Zool. J. Linn. Soc.* **113**, 165 (1995).
4. M. S. Y. Lee, *ibid.* **120**, 197 (1997).
5. _____, *Nature* **379**, 812 (1996).
6. E. S. Gaffney and P. A. Meylan, in *The Phylogeny and Classification of the Tetrapods*, vol. 1, *Amphibians*, *Reptiles*, *Birds*, M. J. Benton, Ed. (Clarendon Press, Oxford, 1988), pp. 157–219.
7. M. Laurin and R. R. Reisz, in *Completing the Transition to Land*, S. S. Sumida and K. L. M. Martin, Eds. (Academic Press, San Diego, CA, 1997), pp. 9–59.
8. O. Rieppel and M. deBraga, *Nature* **384**, 453 (1996).
9. M. deBraga and O. Rieppel, *Zool. J. Linn. Soc.* **120**, 281 (1997).
10. M. Wilkinson, J. Thorley, M. J. Benton, *Nature* **387**, 466 (1997).
11. M. S. Y. Lee, *ibid.* **389**, 245 (1997).
12. W. M. Fitch and E. Margoliash, *Science* **155**, 279 (1967).
13. S. B. Hedges, K. D. Moberg, L. R. Maxson, *Mol. Biol. Evol.* **7**, 607 (1990).
14. K. Fushitani, K. Higashiyama, E. N. Moriyama, K. Imai, K. Hosokawa, *ibid.* **13**, 1039 (1996); J. E. Platz and J. M. Conlon, *Nature* **389**, 246 (1997); H. Mannen, S. C.-M. Tsoi, J. S. Krushkal, W.-H. Li, S. S.-L. Li, *Mol. Biol. Evol.* **14**, 1081 (1997); J. A. W. Kirsch and G. C. Mayer, *Phil. Trans. R. Soc. London Ser. B* **353**, 1221 (1998); T. Gorr, B. K. Mable, T. Kleinschmidt, *J. Mol. Evol.* **47**, 471 (1998). A recent analysis of mitochondrial protein-coding genes [R. Zardoya and A. Meyer, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14226 (1998)] also found an archosaurian affinity of turtles, but no squamates or tuataras were included and the relationships among turtles, birds, and crocodilians could not be resolved.
15. S. B. Hedges, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2621 (1994).
16. Messenger RNAs were prepared from livers of a crocodilian (*Caiman crocodilus*), a tuatara (*Sphenodon punctatus*), a lizard (*Eumeces inexpectatus*), and representatives of the two suborders of turtles (*Trachemys scripta*, a cryptodire, and *Pelusios subniger*, a pleurodire). The coding region of alpha enolase sequenced corresponds to codons 20 to 393 in human. Sequences were generated by means of reverse-transcriptase polymerase chain reaction (PCR) (*15*) and automated DNA sequencing. The 18S rRNA sequence for tuatara was obtained by PCR and automated DNA sequencing. Additional details of the methodology, including primer sequences, are available from the corresponding author. Sequences have been deposited in the GenBank database under accession numbers AF115855-60.
17. The criterion used for selecting genes (13 nuclear and 9 mitochondrial) for phylogenetic analysis was the availability of sequences from at least one turtle, crocodilian, bird, squamate, and mammal. The mammal sequences (rodent and primate) were used for rooting the trees. A subset of those genes (five

nuclear and three mitochondrial) for which a tuatara sequence was available was analyzed separately. An additional 11 nuclear genes were used to estimate divergence times. Genes used for time estimation were selected based on the presence of at least two major groups of reptiles (including birds), at least one mammal sequence for calibration, and a more distant outgroup (*Xenopus* in most cases) for testing rate constancy with respect to the calibration lineage. The mitochondrial data were partitioned into protein-coding (cytochrome b; *cytb*) and non–protein-coding (*mt-npc*; tRNA genes for alanine, asparagine, cysteine, tryptophan, tyrosine, and valine, and the two rRNA genes); these were treated as two separate genes. The 24 nuclear genes are 18*S* ribosomal RNA (*18S rRNA*), 28*S* rRNA, alcohol dehydrogenase (*adh-1*), alpha crystallin A (*a-cryst*), alpha enolase (*a-eno*), alpha globin A chain (*a-glob*), aromatase, beta globin (*b-glob*), calcitonin (*calcit*), calcium-activated potassium channel (*capc*), c-mos proto-oncogene (*c-mos*), cytochrome c (*cytc*), insulin, lactate dehydrogenase A (*ldha*), lactate dehydrogenase B (*ldhb*), lysozyme, myoglobin, pancreatic ribonuclease (*p-ribo*), prolactin, prothrombin (*proth*), rhodopsin, somatotropin (*somato*), superoxide dismutase (*sod*), and tyrosinase (*tyros*). Taxon names and accession numbers are available at www.sciencemag.org/feature/data/986617.shl.

18. Sequence alignments were made with ClustalW [J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acids Res.* **22**, 4673 (1994)] and visually refined. Maximum likelihood, maximum parsimony, and neighbor-joining trees were constructed with MOLPHY [J. Adachi and M. Hasegawa, *MOLPHY Version 2.3: Programs for Molecular Phylogenetics Based on Maximum Likelihood* (*Comp. Sci. Monogr. 28*, Institute of Statistical Mathematics, Tokyo, 1996)], PAUP [D. L. Swofford, *PAUP*. Phylogenetic Analysis using Parsimony* (**and Other Methods*). *Version 4.* (Sinauer Associates, Sunderland, MA, 1998)], and MEGA [S. Kumar, K. Tamura, M. Nei, *MEGA: Molecular Evolutionary Genetic Analysis* (Pennsylvania State University, University Park, PA, 1993)], respectively. Amino acid sequence data were analyzed with the JTT-F model for maximum likelihood and a gamma correction for neighbor-joining. Gene-specific gamma shape parameters were estimated from the amino acid sequence data [X. Gu and J. Zhang, *Mol. Biol. Evol.* **14**, 1106 (1997)]: *A-cryst* (0.33), *A-eno* (0.16), *A-glob* (0.84), *B-glob* (1.05), *Calcit* (>5), *C-mos* (0.46), *Cytb* (0.67), *Cytc* (0.21), *Insulin* (0.82), *Ldha* (0.33), *Ldhb* (0.34), *Myoglobin* (0.82), and combined nuclear proteins (0.38). DNA sequence analyses (neighbor-joining) were performed with a Kimura 2-parameter distance correction (nuclear genes) or Kimura transversion distance correction (mitochondrial genes). Statistical confidence estimates of topologies and branches were obtained with the RELL bootstrap method for maximum likelihood, the standard bootstrap method (5000 replications) for maximum parsimony and neighbor-joining, and with the four-cluster and interior-branch tests [S. Kumar, *Phyltest: A Program for Testing Phylogenetic Hypotheses* (Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA, ed. 2.0, 1996)]. Alignments are available from the corresponding author.
19. A. Rzhetsky, S. Kumar, M. Nei, *Mol. Biol. Evol.* **12**, 163 (1995).
20. Eleven of the original 13 nuclear genes, the two mitochondrial data partitions, and 11 additional nuclear genes were used for estimating pairwise divergence times [S. B. Hedges, P. H. Parker, C. G. Sibley, S. Kumar, *Nature* **381**, 226 (1996); S. Kumar and S. B. Hedges, *ibid.* **392**, 917 (1998)] among the major groups of reptiles. Two genes, calcitonin and insulin, were omitted because of their short (<100 amino acids) sequence lengths. The fossil divergence between mammals and reptiles (including birds) at 310 Ma was used as the calibration. For each divergence time estimate, three rate constancy tests [N. Takezaki, A. Rzhetsky, M. Nei, *Mol. Biol. Evol.* **12**, 823 (1995)] were performed: the two lineages in question and each lineage versus the calibration lineage. Divergence times were used only if rate constancy was not rejected in all three tests. These constant-rate divergence times were then averaged between groups defined in the phylogenetic analysis of combined data (Fig. 1B), and their standard errors were estimated. The upper and lower 5% (or at least the highest and lowest) time estimates were excluded to minimize the effect of outliers caused by gene paralogy or other biases. Additional details are available at www.sciencemag.org/feature/data/986617.shl.
21. M. J. Benton, *The Fossil Record 2* (Chapman & Hall, London, 1993).
22. M. Goodman, M. M. Miyamoto, J. Czelusniak, in *Molecules and Morphology in Evolution: Conflict or Compromise?* C. Patterson, Ed. (Cambridge Univ. Press, Cambridge, 1987); M. J. Bishop and A. E. Friday, in (*6*), pp. 33–58.
23. A. Saita, M. Comazzi, E. Perrotta, *Boll. Zool.* **4**, 307 (1987).
24. B. G. M. Jamieson and J. W. Healy, *Philos. Trans. R. Soc. London Ser. B* **335**, 207 (1992).
25. A. S. Romer, *Osteology of the Reptiles* (Univ. of Chicago Press, Chicago, IL, 1956).
26. E. S. Gaffney, *Bull. Am. Mus. Nat. Hist.* **194**, 1 (1990).
27. We thank S. Pyott and M. van Tuinen for laboratory assistance; X. Gu and J. Zhang for the program to estimate the gamma parameter; C. Hass, S. Kumar, M. van Tuinen, and R. Thomas for comments and discussion; and C. Daugherty and L. Maxson for materials. This research was supported in part by a grant from the Howard Hughes Medical Institute (Undergraduate Biological Sciences Education Program) to Penn State and by grants from the NSF (DEB-9615643) and NASA (NCC2-1057) to S.B.H.

19 November 1998; accepted 6 January 1999

Efficient Bypass of a Thymine-Thymine Dimer by Yeast DNA Polymerase, Polη

Robert E. Johnson, Satya Prakash, Louise Prakash*

The *RAD30* gene of the yeast *Saccharomyces cerevisiae* is required for the error-free postreplicational repair of DNA that has been damaged by ultraviolet irradiation. Here, *RAD30* is shown to encode a DNA polymerase that can replicate efficiently past a thymine-thymine *cis-syn* cyclobutane dimer, a lesion that normally blocks DNA polymerases. When incubated in vitro with all four nucleotides, Rad30 incorporates two adenines opposite the thymine-thymine dimer. Rad30 is the seventh eukaryotic DNA polymerase to be described and hence is named DNA polymerase η.

Cells possess a variety of mechanisms to repair damaged DNA. If left unrepaired, DNA lesions in the template strand can block the replication machinery (*1*). In *S. cerevisiae*, genes in the *RAD6* epistasis group function in the replication of DNA that has been damaged by ultraviolet (UV) light and other agents (*2*). Of the genes in this group, *REV1*, *REV3*, and *REV7* are required for UV mutagenesis, whereas *RAD5* and *RAD30* function in alternate pathways of error-free bypass of UV-induced DNA damage (*3, 4*). Rev3 and Rev7 associate to form a DNA polymerase, Polζ, that can weakly bypass *cis-syn* thymine-thymine (T-T) dimers (*5*). Rev1 has a deoxycytidyl transferase activity that transfers a deoxycytidine 5′-monophosphate residue to the 3′ end of a DNA primer in a template-dependent reaction (*6*). The addition of Rev1 has no effect on the Polζ bypass of *cis-syn* T-T dimers (*6*). Rad5 is a DNA-dependent adenosine triphosphatase (*7*) and Rad30 is homologous with *Escherichia coli* DinB and UmuC and with *S. cerevisiae* Rev1 (*4*). Here, we elucidate the function of yeast Rad30.

The *RAD30* gene was fused in-frame downstream of the glutathione S-transferase (*GST*) gene, and the resulting fusion protein (~95 kD) was purified to near homogeneity (*8*) from a protease-deficient yeast strain harboring the GST-Rad30 expression plasmid pBJ590. Plasmid pBJ590 fully complements the UV sensitivity of the *rad30*Δ mutation, indicating that the Rad30 fusion protein functions normally in vivo. We first examined whether Rad30 has deoxynucleotidyl transferase activity (*9*). DNA substrates containing a different template nucleotide at the primer-template junction were incubated with Rad30 in the presence of just one deoxynucleoside triphosphate (dNTP). Rad30 predominantly incorporated the correct nucleotide across from each of the four template nucleotides, but incorrect nucleotides were also inserted across from some template residues (Fig. 1A). For instance, a C residue and a T residue were weakly incorporated opposite template C (Fig. 1A, lane 6) and opposite template G (Fig. 1A, lane 11), respectively. Rad30 exhibited the lowest specificity when the template nucleotide was T, as in substrate S-3. Whereas the A residue was incorporated with the highest efficiency (Fig. 1A, lane 16), G, T, and C residues were also incorporated with varying efficiencies (Fig. 1A, lanes 15, 17, and 18, respectively). The pattern of misincorporations seemed to depend on the sequence context of the template. For example,

Sealy Center for Molecular Science, University of Texas Medical Branch at Galveston, 6.104 Medical Research Building, 11th and Mechanic Streets, Galveston, TX 77555–1061, USA.

*To whom correspondence should be addressed. E-mail: lprakash@scms.utmb.edu