Volume 334, number 1, 27–31 FEBS 13201 November 1993

Molecular characterization of a blue visual pigment gene in the fish *Astyanax fasciatus*

Ruth Yokoyama, Shozo Yokoyama*

Department of Biology, Syracuse University, 130 College Place, Syracuse, NY 13244, USA

Received 23 August 1993; revised version received 17 September 1993

We report here the isolation and sequence determination of a gene closely linked to the *Astyanax* red visual pigment gene. Reverse transcription polymerase chain reaction assays show that this new gene ($B23_{Af}$) and the previously characterized red and green visual pigment genes of *Astyanax* are all expressed in the eye. Phylogenetic analysis shows that $B23_{Af}$ belongs to the group consisting of short wavelength-sensitive pigment genes from different species and is most closely related to the goldfish blue visual pigment gene.

Blue visual pigment gene; Opsin; Molecular evolution; *Astyanax fasciatus*

1. INTRODUCTION

The retina of the eyed Mexican characin, *Astyanax fasciatus*, contains rods and single and double cones. The rods function in dim light and do not perceive color. The single cones are responsible for blue vision, whereas one member of the double cone is red-sensitive and the other is green-sensitive [1]. The red and green visual pigment genes from a blind cavefish of the same species, *A. fasciatus* ([2]; references therein), have previously been characterized [3–5]. The λ clone containing the *Astyanax* red visual pigment gene cross-hybridized to the bovine rhodopsin and human blue cDNA probes. A DNA segment 6 kilobases (kb) upstream from the red gene was found to have high sequence similarities to visual pigment genes. We report here the cloning and sequencing of this gene (named $B23_{Af}$). Phylogenetic and expression analyses of this and other visual pigment genes strongly suggest that $B23_{Af}$ is a blue visual pigment gene.

2. MATERIALS AND METHODS

2.1. λ clone characterization and DNA sequencing

To characterize the visual pigment gene that is physically linked to $R007_{Af}$, a red visual pigment gene from clone λ007 [4,5], an overlapping clone λ23 was isolated from the same genomic library (see [3]). The maps of λ007 and λ23 are shown in Fig. 1. Subcloning and double stranded sequencing were conducted as previously described [3,4].

2.2. RNA isolation and quantitative RT/PCR amplification

Total RNA from whole eyes of 24-day-old *Astyanax fasciatus* was isolated by a guanidinium isothiocyanate procedure [6].

A quantitative performance of the reverse transcription polymerase chain reaction (RT/PCR) assay was determined by initially trying various RNA dilutions with 20, 25 and 30 amplification cycles. 25 cycles of amplification showed a linear increase of PCR products within the 20-fold range of RNA concentration used. RNA was mixed with the PCR reaction mix (10 mM Tris-HCl, pH 9.0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, 200 μM dNTPs and primers at 1 μM each), and ribonuclease inhibitor, Taq polymerase and Moloney murine leukemia virus reverse transcriptase. The samples were placed in a thermal cycler at 50°C for 8 min, followed by 25 cycles of 94°C for 1 min, 60°C for 2 min and 72°C for 2.5 min. The following four sets of primers were used: (1) **007**: 5′-TTTCTTTGCCTGCTTTGCGGC-GGC-3′(F); 5′-GTGTGAGGTATACATGGTGTTTTA-3′ (R) (positions 1786 and 2134 of $R007_{Af}$, [5]); (2) **23**: 5′-GGGTGGTGTTTA-ACCGCGGGCAAT-3′ (F); 5′-CCGTTAGATTTATTTCTCCGGT-CC-3′ (R) (codons 279 to 9 nucleotides 3′-flanking sequence of $B23_{Af}$, this paper); (3) **23′**: 5′-CCAGACTGGTACACCACTGAAAAC-3′ (F); 5′-CGTTCCCAACCGCAGATCAAACGA-3′ (R) (codons 194 through 294 of $B23_{Af}$, this paper), and (4) **103**: 5′-CCTGGCACCC-ACTGGCAGCTGCTC-3′ (F); 5′-AGTAAAGGGCGCCATAGAT-TTTTA-3′ (R) (codon 295 to 21 nucleotides 3′-flanking sequence of $G103_{Af}$, [3]).

2.3. Data analyses

Using invertebrate rhodopsin genes as the outgroup, the rooted phylogenetic tree of 27 currently known vertebrate visual pigment genes was first constructed (results not shown). The evolutionary tree obtained shows that (1) $B23_{Af}$ is evolutionarily closely related to the visual pigment genes from human (B_{Hs}; [7]), chicken (B_{Gg} and V_{Gg}; [8]), and goldfish (B_{Ca}; [9]) and (2) the UV-absorbing visual pigment gene from zebrafish (UV_{Br}; [10]) is most closely related to the rhodopsin gene from goldfish (Rh_{Ca}; [9]).

To consider all known visual pigment genes from fishes in the analyses, two sets of visual pigment genes were considered: (1) the blue and closely related short wavelength-sensitive (SWS) genes ($B23_{Af}$, B_{Hs}, V_{Gg}, B_{Ca}, and B_{Gg}); and (2) rhodopsin and related genes. For the second group, the rhodopsin genes were selected from lamprey (Rh_{Lf}, [11]), chicken (Rh_{Gg}; [12]), human (Rh_{Hs}; [13]), Rh_{Ca} from goldfish, and UV_{Br} from zebrafish as representatives.

To evaluate the evolutionary relationship of these 10 visual pigment genes, the amino acid sequences deduced from them were subjected to statistical analysis. The proteins encoded by the visual pigment genes B_{Hs}, V_{Gg}, B_{Ca}, $B23_{Af}$, B_{Gg}, Rh_{Lf}, Rh_{Ca}, Uv_{Br}, Rh_{Gg}, and Rh_{Hs} were

*Corresponding author. Fax: (1) (315) 443 2012.
Email: syokoyam@suvm.acs.syr.edu. Bitnet: syokoyam@suvm.

denoted as B_{Hs}, V_{Gg}, B_{Ca}, $B23_{Af}$, B_{Gg}, Rh_{Lj}, Rh_{Ca}, UV_{Bt}, Rh_{Gg}, and Rh_{Hs}, respectively. The number of amino acid replacements per site for two polypeptides (K) was estimated by $K = -\ln(1-p)$, where p is the proportion of different amino acids between the two sequences. Topology and branch lengths of the phylogenetic tree were evaluated by applying the neighbor-joining (NJ) method [14] to the K values. Bootstrap probabilities for branches of the NJ tree for the amino acid replacements were estimated by bootstrap analysis with 1000 replications (CLUSTAL V; [15]).

3. RESULTS AND DISCUSSION

3.1. *Nucleotide sequence of* $B23_{Af}$

The sequencing strategy and nucleotide sequence of $B23_{Af}$ are shown in Figs. 1 and 2, respectively. This gene has 5 exons and 4 introns, as is the case for the vertebrate rhodopsin genes [12,13,16] and B_{Hs} [7]. This is unlike the red- and green-sensitive genes which have 6 exons [3,4,7]. The GT/AG consensus intron splice junctions are conserved. Although $B23_{Af}$ is only about 6 kb away from $R007_{Af}$, their sequences differ considerably. For example, when the sequence of the coding region of $B23_{Af}$ was compared to those of the red- and green-sensitive visual pigment genes in human [7] and fish [3–5], human blue visual pigment gene [7] and rhodopsin genes from lamprey [11], human [13], and chicken [12], the proportions of identical nucleotides are 51–54%, 57%, and 58–62%, respectively. However, $B23_{Af}$ is about 74% identical at the nucleotide level to the goldfish blue visual pigment gene [9].

$B23_{Af}$ shares several structural features with other opsins: (1) a lysine at position 301, important for Schiff base linkage to a chromophore [17,18], (2) a glutamic acid at position 118, the Schiff base counterion [19–22], (3) cysteines at residues of 115 and 192, the sites for an essential disulfide bond [23], (4) cysteine at 145, involved in phosphorylation through interacting with the C-terminal tail [24], (5) glutamic acid and arginine at 139 and 140, respectively, believed to be important for transducin interaction [25], and (6) C-terminal serines and threonines, potential sites for phosphorylation [26]. It is interesting that $B23_{Af}$ position 297 is serine like the blue opsins in goldfish [9], human [7], and chicken [8], whereas all other opsins contain alanine at this position and may be important for interaction with retinal [9].

3.2. *Expression of the color visual pigments*

To determine $B23_{Af}$ is a functional gene, primers were designed to reverse transcribe and PCR amplify 2 exons of the red ($R007_{Af}$), green ($G103_{Af}$) and $B23_{Af}$. Conditions were chosen for a linear increase dependent upon the RNA concentration to avoid plateau amplification of rare RNAs (see section 2). Fig. 3 shows that all three of these genes are expressed in the eye in approximately equivalent amounts. The PCR products were subcloned and/or directly sequenced to verify that each band consisted of the expected visual pigment (data not shown).

3.3. *A phylogenetic tree*

The phylogenetic tree for the SWS visual pigments and rhodopsins is shown in Fig. 4. The identical tree topology was obtained when the NJ method was applied to K values for all known visual pigments (result not shown). The same result was also obtained when the NJ method was applied to the numbers of nucleotide substitutions per site for pairwise comparison, which was estimated by using Kimura's [27] two-parameter method (result not shown). The bootstrap analysis supports that the distinction between the SWS visual pigments (SWS group) and the rhodopsins (Rhodopsin group) is highly reliable. Note that $B23_{Af}$ is most closely related to B_{Ca}. Within the SWS group, B_{Hs} and V_{Gg} form one subgroup and B_{Ca}, $B23_{Af}$ and B_{Gg} form another.

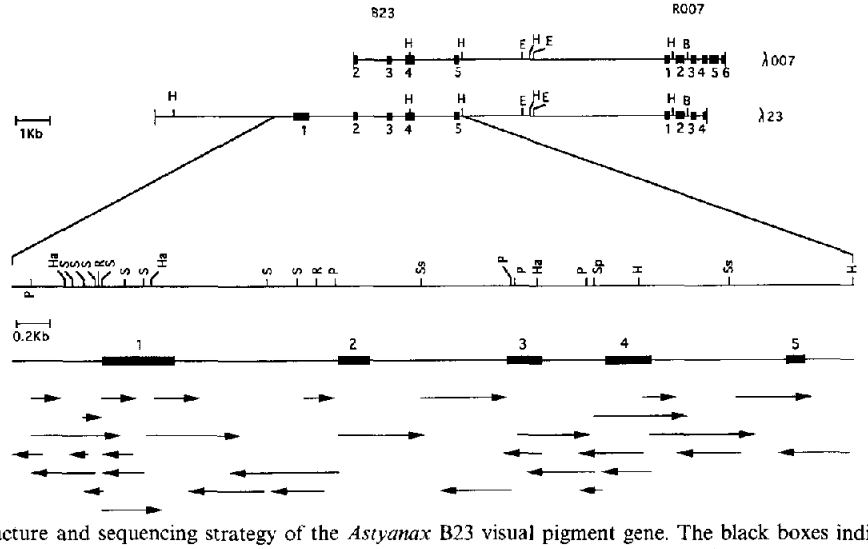

Fig. 1. The genomic structure and sequencing strategy of the *Astyanax* B23 visual pigment gene. The black boxes indicate the putative coding sequence. The genomic clone was subcloned into Bluescript SK⁻ plasmid by using the indicated restriction enzyme or exonuclease III/mung bean nuclease deletions. (B) *Bam*HI; (E) *Eco*RI; (H) *Hin*dIII; (Ha) *Hae*III; (P) *Pst*I; (R) *Rsa*I; (S) *Sau*3A I; (Sp) *Spe*I; and (Ss) *Sst*I.

```
                                                  CCT GCAGCAAAGC AGATTAAAGA GGAATGTCTT ATACAATTCA ATTCTACCTG
TGTGTCTTTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGAGTGTG TGTGTGTGTG TGTGTAAGAG AGAGAGAGAG AGTATTAGTC AGTAATTAAT GGGGACGACC
CTATTAGGCC CCTATAAAAG ATGACAGGAG ATCCTACAGC TCAGAGATCA GGGTCAGTAA CCGGTCGGTC CCCCGTCTCC TGTCTCCTGT CCTCTGTCTC CTGTCTCACC
GCTGATCTGA TGAGGTTCCA GCTCTCGGTA ATCGTGTAGA TTTCTGTAAG AACTCTAAAG GAACCGGTTC TGCTGATCAC TGTACTGTAG ATCATCATCC CTGAGTCAGA
                             10                                      20                                      30
Met Lys Ser Arg Pro Gln Glu Phe Gln Glu Asp Phe Tyr Ile Pro Ile Pro Leu Asp Thr Asn Asn Ile Thr Ala Leu Ser Pro Phe Leu
ATG AAG AGT CGT CCG CAG GAG TTT CAG GAG GAT TTC TAC ATC CCC ATC CCT CTG GAC ACC AAT AAC ATC ACG GCC CTC AGC CCG TTC CTG
                             40                                      50                                      60
Val Pro Gln Asp His Leu Gly Gly Ser Gly Ile Phe Met Ile Met Thr Val Phe Met Leu Phe Leu Phe Ile Gly Gly Thr Ser Ile Asn
GTC CCG CAG GAC CAT CTA GGA GGG TCA GGG ATT TTC ATG ATC ATG ACC GTC TTC ATG CTT TTC CTG TTT ATT GGT GGA ACC AGC ATC AAC
                             70                                      80                                      90
Val Leu Thr Ile Val Cys Thr Val Gln Tyr Lys Lys Leu Arg Ser His Leu Asn Tyr Ile Leu Val Asn Leu Ala Ile Ser Asn Leu Leu
GTC CTC ACC ATC GTC TGC ACC GTC CAG TAC AAG AAG CTC CGA TCA CAT CTC AAC TAC ATC CTG GTG AAC CTG GCC ATC TCC AAC CTG CTG
                            100                                     110                                     120
Val Ser Thr Val Gly Ser Phe Thr Ala Phe Val Ser Phe Leu Asn Arg Tyr Phe Ile Phe Gly Pro Thr Ala Cys Lys Ile Glu Gly Phe
GTC TCC ACC GTC GGG TCC TTC ACC GCC TTC GTC TCC TTC CTC AAC CGA TAC TTT ATT TTT GGA CCC ACA GCT TGT AAA ATT GAG GGA TTT
                                                                 130                                 140
Val Ala Thr Leu Gly G                            ly Met Val Ser Leu Trp Ser Leu Ser Val Val Ala Phe Glu Arg Trp Leu
GTT GCA ACT TTA GGA G  GTAATGTGTT  867bp  CTCTCTGCAG  GT ATG GTG AGT CTG TGG TCT CTG TCG GTC GTG GCG TTT GAG AGG TGG CTG
                     150                                     160                                     170
Val Ile Cys Lys Pro Val Gly Asn Phe Ser Phe Lys Gly Thr His Ala Ile Ile Gly Cys Ala Leu Thr Trp Phe Phe Ala Leu Leu Ala
GTT ATC TGT AAA CCT GTG GGG AAT TTC TCC TTT AAA GGA ACT CAC GCT ATA ATC GGC TGT GCT CTC ACC TGG TTT TTC GCT CTG CTC GCC
                     180                                                                         190
Ser Thr Pro Pro Leu Phe Gly Trp Ser Ar                                     g Tyr Ile Pro Glu Gly Leu Gln Cys Ser Cys Gly Pro
TCC ACA CCT CCA CTG TTC GGC TGG AGC AG  GTCAGATATA  765bp  TTTTCTGCAG  G TAC ATC CCT GAA GGT CTG CAG TGT TCC TGT GGT CCA
                 200                                      210                                      220
Asp Trp Tyr Thr Thr Glu Asn Lys Tyr Asn Asn Glu Ser Tyr Val Met Phe Leu Phe Cys Phe Cys Phe Gly Phe Pro Phe Thr Val Ile
GAC TGG TAC ACC ACT GAA AAC AAA TAC AAC AAC GAG TCT TAT GTC ATG TTC CTC TTC TGC TTC TGC TTC GGA TTC CCG TTT ACT GTT ATC
                 230                                                                       240
Leu Phe Cys Tyr Gly Gln Leu Leu Phe Thr Leu Lys Ser                                        Ala Ala Lys Ala Gln Ala Asp Ser Ala
CTT TTC TGC TAC GGC CAA CTG CTC TTC ACT CTC AAA TCA  GTGAGACACC  325bp  TGTCCACTAG  GCG GCC AAA GCT CAG GCT GAC TCC GCC
         250                                     260                                     270
Ser Thr Gln Lys Ala Glu Arg Glu Val Thr Lys Met Val Val Val Met Val Met Gly Phe Leu Val Cys Trp Leu Pro Tyr Ala Ser Phe
TCC ACG CAG AAG GCG GAG CGA GAG GTG ACT AAG ATG GTG GTG GTG ATG GTG ATG GGG TTT CTG GTG TGC TGG CTG CCC TAC GCC TCC TTC
         280                                     290                                     300
Ala Leu Trp Val Val Phe Asn Arg Gly Gln Ser Phe Asp Leu Arg Leu Gly Thr Ile Pro Ser Cys Phe Ser Lys Ala Ser Thr Val Tyr
GCC CTG TGG GTG GTG TTT AAC CGC GGG CAA TCG TTT GAT CTG CGG TTG GGA ACG ATA CCG TCC TGC TTC TCT AAA GCT TCT ACC GTC TAC
         310                                                                             320
Asn Pro Val Ile Tyr Val Phe Met Asn Lys Gln                                     Phe Arg Ser Cys Met Met Lys Leu Ile Phe Cys
AAT CCC GTA ATC TAT GTC TTC ATG AAC AAA CAG  GTGCAGCAAT  754bp  TTGTGTGCAG  TTC CGC TCC TGC ATG ATG AAG CTG ATT TTC TGT
     330                                     340                                     350
Gly Lys Ser Pro Phe Gly Asp Asp Glu Glu Ala Ser Ser Ser Ser Gln Val Thr Gln Val Ser Ser Val Gly Pro Glu Lys
GGG AAG AGT CCG TTT GGA GAT GAT GAA GAA GCC TCC TCC TCC TCT CAG GTG ACC CAG GTG TCT TCT GTA GGA CCG GAG AAA TAA

ATCTAACGGA CAGAAGAATC ACAAACACCT CCAGCGTCAC ACAGCGTCAT AAAGCCTCGG GAAGTCTGTT AGACGTTTGC AGGAAGGTGC GTCTCTGAGG CTCAGGTTAT
TCAGCAGAGC ACGAACCGTC ATACTCCCTC CTTTTTCCCT CTGAGAACAT TCTGCTTTTC CTCTCTGATC
```

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the *Astyanax* $B23_{Af}$ gene. The deduced amino acid sequence is written above each nucleotide triplet.

Distinction of these two subgroups is again well supported by the bootstrap analysis (see Fig. 4). The two subgroups also reflect some functional differences. That is, the visual pigments B_{Hs} and V_{Gg} have absorption maxima of 424 nm [28] and 415 nm [8], respectively, whereas B_{Gg} and B_{Ca} have much longer absorption maxima at 455 nm [8] and 441 nm [9], respectively. The divergence of the ancestral genes of B_{Gg} and V_{Gg} in chicken predates the speciation event between fish and chicken. Clearly, the gene duplication between the two groups of SWS genes preceded the divergence between fish and bird, some 400 million years ago.

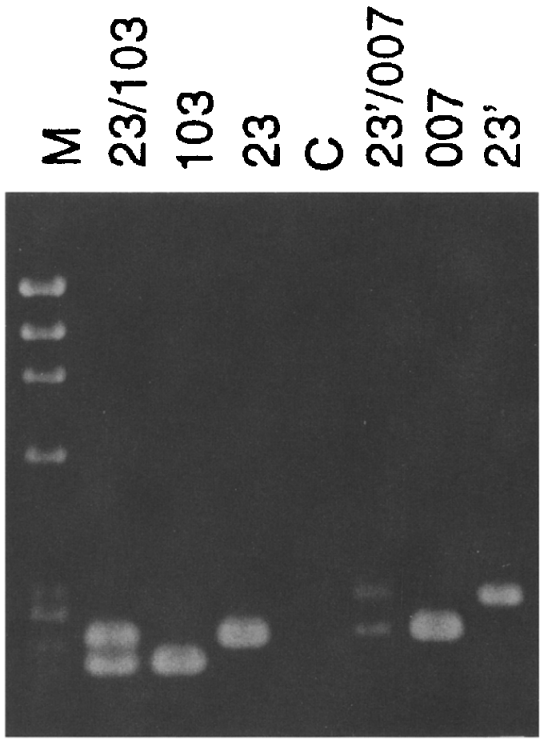

Fig. 3. Agarose gel stained with ethidium bromide, demonstrating expression of *Astyanax* visual pigment genes. RNA isolated from 24-day-old fish eyes were reverse transcribed and PCR amplified with primers designed to amplify the last two exons of $B23_{Af}$ (lanes labelled **23**), $G103_{Af}$ (lanes labelled **103**), and $R007_{Af}$ (lanes labelled **007**), and exons 3 and 4 of $B23_{Af}$ (labelled **23'**). Lanes marked 23/103 and 23'/007 had both sets of primers indicated in the RT/PCR reaction. Lane M contains ϕXHaeIII as a marker. Lane C contains the same PCR sample as lane 2 without reverse transcriptase to verify that amplification requires reverse transcription.

As noted before, four different types of visual pigments are distinguished in *Astyanax*, i.e. rhodopsin, blue-, green-, and red-sensitive visual pigments [1]. The red and green visual pigments belong to a different group which is distantly related to the SWS group [8,9] and an *Astyanax* rhodopsin is closely related to Rh_{Ca} and UV_{Br} (Yokoyama and Yokoyama, unpublished). Since $B23_{Af}$ is expressed in the *Astyanax* retina and has high homology to the goldfish blue, it is highly likely that this gene encodes a blue opsin.

The ancestral genes encoding the human red (and green) and blue visual pigments and rhodopsins seem to have arisen more than 500 million years ago [29,30]. Because of their long evolutionary histories, it is not surprising to see that the three major groups of visual pigment genes are localized in different chromosomes. It turns out that, in human, the rhodopsin and the blue visual pigment genes are located on chromosomes 3 and 7, respectively, and the tandemly located red and green visual pigment genes are on X-chromosomes [29]. Since human and fish diverged about 400 million years ago, the three groups of visual pigment genes in *Astyanax* seem to share the same three different common ancestors of the human visual pigment genes [30]. It is therefore totally unexpected to see that $B23_{Af}$ and $R007_{Af}$ are still linked with each other so closely.

Acknowledgements: This study was supported by USPHS Grant GM-42379.

REFERENCES

[1] Kleinschmidt, J. and Harosi, F.I. (1992) Proc. Natl. Acad. Sci. USA 89, 9181 9185.
[2] Wilkens, H. (1988) Evol. Biol. 23, 271 367.
[3] Yokoyama, R. and Yokoyama, S. (1990) Vision Res. 30, 807–816.
[4] Yokoyama, R. and Yokoyama, S. (1990) Proc. Natl. Acad. Sci. USA 87, 9315–9318.
[5] Yokoyama, S., Starmer, W.T. and Yokoyama, R. (1993) Mol. Biol. Evol. 10, 527–538.
[6] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.

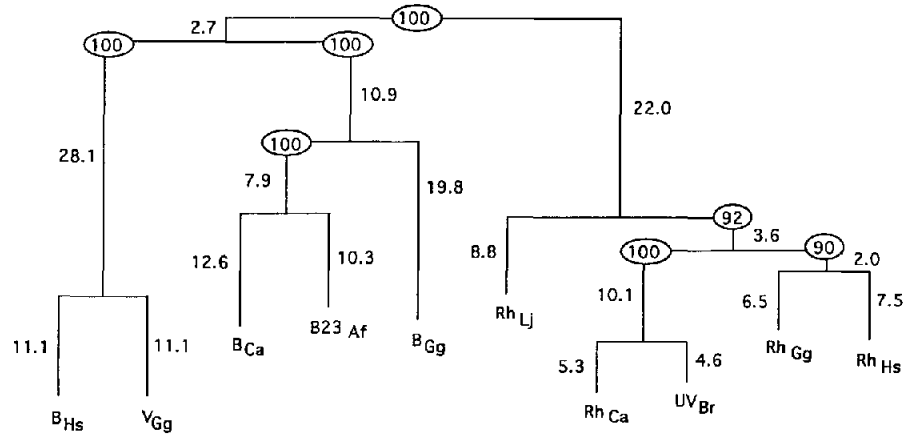

Fig. 4. Unrooted phylogenetic tree for the blue visual pigments and rhodopsins. The numbers next to the different branches are the numbers of amino acid replacements (× 100), as estimated by means of the NJ method [14]. Circled numbers indicate clustering percent support generated by 1000 bootstrap analyses [15].

[7] Nathans, J., Thomas, D. and Hogness, D.S. (1986) Science 232, 193–202.
[8] Okano, T., Kojima, D., Fukada, Y., Shichida, Y. and Yoshizawa, T. (1992) Proc. Natl. Acad. Sci. USA 89, 5932–5936.
[9] Johnson, R., Grant, K.B., Zankel, T.C., Boehm, M.F., Merbs, S.L., Nathans, J. and Nakanishi, K. (1993) Biochemistry 32, 208–214.
[10] Robinson, J., Schmitt, E.A., Harosi, F., Reece, R.J. and Dowling, J.E. (1993) Proc. Natl. Acad. Sci. USA 90, 6009–6012.
[11] Hisatomi, O., Iwasa, T., Tokunaga, F. and Yasui, A. (1991) Biomed. Biophys. Res. Commun. 174, 1125 1132.
[12] Takao, M., Yasui, A. and Tokunaga, F. (1988) Vision Res. 28, 471 480.
[13] Nathans, J. and Hogness, D.S. (1984) Proc. Natl. Acad. Sci. USA 81, 4851–4855.
[14] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[15] Higgins, D.G., Bleasby, A.J. and Fuchs, R. (1992) Comput. Appl. Biosci. 8, 189–191.
[16] Nathans, J. and Hogness, D.S. (1983) Cell 34, 807–814.
[17] Bownds, D. (1967) Nature 216, 1178–1181.
[18] Wang, J.K., McDowell, H. and Hargrave, P.A. (1980) Biochemistry 19, 5111–5117.
[19] Zhokovsky, E.A. and Oprian, D.D. (1989) Science 246, 928–930.
[20] Sakmar, T.P., Franke, R.R. and Khorana, H.G. (1989) Proc. Natl. Acad. Sci. USA 86, 8309–8313.
[21] Nathans, J. (1990) Biochemistry 29, 9746–9752.
[22] Sakmar, T.P., Franke, R.R. and Khorana, H.G. (1991) Proc. Natl. Acad. Sci. USA 88, 3079–3083.
[23] Karnik, S.S. and Khorana, H.G. (1990) J. Biol. Chem. 265, 17520–17524.
[24] Karnik, S.S., Ridge, K.D., Bhattacharya, S. and Khorana, H.G. (1993) Proc. Natl. Acad. Sci. USA 90, 40–44.
[25] Franke, R.R., Konig, B., Sakmar, T.P., Khorana, H.G. and Hofmann, K.P. (1990) Science 250, 123–125.
[26] Palczewski, K., McDowell, J.H. and Hargrave, P.A. (1988) J. Biol. Chem. 263, 14067 14073.
[27] Kimura, M. (1980) J. Mol. Evol. 16, 111 120.
[28] Oprian, D.D., Asenjo, A.B., Lee, N. and Pelletier, S.L. (1991) Biochemistry 30, 11367–11372.
[29] Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T.B. and Hogness, D.S. (1986) Science 232, 203–210.
[30] Yokoyama, S. and Yokoyama, R. (1989) Mol. Biol. Evol. 6, 186–197.