Hisao Ueyama · Ryuzo Torii · Shoko Tanabe
Sanae Oda · Shinichi Yamade

An insertion/deletion *TEX28* polymorphism and its application to analysis of red/green visual pigment gene arrays

Received: 21 June 2004 / Accepted: 22 July 2004 / Published online: 17 September 2004
© The Japan Society of Human Genetics and Springer-Verlag 2004

Abstract *TEX28* gene (fTEX) is present immediately downstream of the red/green visual pigment gene array on the human X chromosome. Its pseudogene (pTEX) that lacks exon 1 is present within the array between pigment genes. We found that both fTEX and pTEX genes had a 697 bp insertion/deletion polymorphism in their introns 3. In color-normal male subjects, the frequency of the 697 bp region was 43% (40/94) in pTEX and 97% (91/94) in fTEX in the array of Red-pTEX-Green-fTEX and 10% (9/94) in pTEX and 87% (41/47) in fTEX in the array of Red-pTEX-Green-pTEX-Green-fTEX. These results suggest that normal arrays with multiple green genes may have arisen through gene duplication rather than unequal homologous crossover. In color-vision-deficient male subjects with a single-gene array, the frequency of the 697 bp region was 83% (25/30) in the array of Green-fTEX and 66% (74/112) in the array of Red-fTEX. In color-vision-deficient male subjects with a 2-gene array, the frequency of the region was 44% (16/36) in pTEX and 97% (35/36) in fTEX in the array of Green-pTEX-Green-fTEX and 75% (18/24) in pTEX and 92% (22/24) in fTEX in the array of Red-pTEX-Red-fTEX. These results suggest that 2-green-gene arrays have arisen through unequal

homologous crossover between a normal 2-gene array and a single-green-gene array. With data from a long-range PCR method using the insertion/deletion polymorphism, we proposed a structure of the second gene of 3-gene arrays, Green-pTEX-Green-pTEX-Green-fTEX and Red-pTEX-Red-pTEX-Red-fTEX, in color-vision-deficient subjects.

Keywords Visual pigment gene · Long-range PCR · *TEX28* · Congenital color-vision deficiency

Introduction

The red/green cone visual pigment genes that are present in tandem on the long arm of the human X chromosome constitute the visual pigment gene array (Nathans et al. 1986a). In the arrays of color-normal males, a single red gene is followed by one or more green genes. Gross rearrangement of the array leads to congenital color-vision deficiencies. A 5'-red-green hybrid gene is associated with the protan color-vision deficiency that involves the loss of the red pigment and a 5'-green-red hybrid gene or a deletion of a green gene is associated with the deutan color-vision deficiency that involves the loss of the green pigment (Nathans et al. 1986b; Deeb et al. 1992). The red and green pigments differ by only 15 amino acids, but the most important differences are at the amino acid residues at 277 and 285. They are 277Tyr and 285Thr in the red pigment and 277Phe and 285Ala in the green pigment. These amino acid changes result in a shift of $\lambda_{\max}$ by 20–23 nm, which accounts for most of the difference in $\lambda_{\max}$ between the red and green pigments (Merbs and Nathans 1992; Asenjo et al. 1994). Since both amino acid residues are encoded by exon 5, in this study, we discriminated between red and green pigment genes with the differences in exon 5.

The red/green pigment genes are thought to have evolved from their ancestral gene through gene duplication because most mammals have only one type of visual pigment for middle-wave-length to long-wave-length

H. Ueyama (✉)
Department of Medical Biochemistry,
Shiga University of Medical Science,
Seta, Otsu 520-2192, Japan
E-mail: datt@belle.shiga-med.ac.jp
Tel.: +81-77-5482163
Fax: +81-77-5482164

R. Torii
Research Center for Animal Life Science,
Shiga University of Medical Science,
Otsu, Japan

S. Tanabe
Institute of Visual Science, Nagoya, Japan

S. Oda · S. Yamade
Department of Ophthalmology,
Shiga University of Medical Science,
Otsu, Japan

lights (green–yellow–red region) (Nei et al. 1997). Only humans, apes, old-world monkeys, and howler monkeys, which are one of the new-world monkeys, have two visual pigments for middle-wave-length to long-wave-length lights. The *TEX28*, or *CXorf2*, gene, which had been present immediately downstream of the ancestral gene, was duplicated together, but its exon 1 was excluded from the process (Hanna et al. 1997). Therefore, the *TEX28* gene between visual pigment genes lacks exon 1 and is thought to be nonfunctional. We call this intergenic nonfunctional gene pTEX in this study and the original functional gene fTEX. On comparison of nucleotide sequences between fTEX and pTEX genes, we found a 697 bp insertion/deletion polymorphism in their introns 3. In this communication, we present the frequency of the polymorphism in color-normal and color-vision-deficient Japanese male subjects.

The red/green visual pigment gene array usually consists of multiple genes. The first gene of the array can be completely amplified in vitro by the long-range PCR method using a first gene-specific primer (Oda et al. 2003). However, downstream genes of the array can be completely amplified all together (Oda et al. 2003) because they have either the same or very similar nucleotide sequences, including their upstream and downstream regions. Therefore, individual analysis of each of the downstream genes in a 3-or-more gene array has not been possible. Since there is good evidence for the hypothesis that only the first two genes of the array are expressed in the human retina (Hayashi et al. 1999), information about a structure of the second gene of the array may be necessary for prediction of color-vision phenotypes by gene analysis. In this communication, we describe usefulness of the insertion/deletion polymorphism as it has been applied to the analysis of 3-gene arrays.

Materials and methods

Subjects

Genomic DNA was extracted from the peripheral blood leukocytes of volunteer color-normal Japanese males.

Their color-vision status was assessed using the Ishihara pseudoisochromatic plates. Genomic DNA was also extracted from the peripheral blood leukocytes of color-vision-deficient Japanese males who had consulted Shiga University of Medical Science Hospital or Japan Red Cross Nagoya First Hospital. Their color-vision status was assessed by clinical tests including examinations using a Nagel anomaloscope (model I, Schmidt & Haensch, Berlin). They were informed of the aims and methods of this study and gave their consent. All procedures were conducted in accordance with the declaration of Helsinki. Genomic DNA was also extracted from the peripheral blood leukocytes of two female monkeys—a Japanese monkey (*Macaca fuscata*) and a crab-eating monkey (*Macaca fascicularis*)—which had been housed individually and maintained at the Research Center for Animal Life Science, Shiga University of Medical Science.

Gene analysis

The number of red/green visual pigment genes was estimated from the ratio of PCR products for promoters using primers PF-2 and PR (Table 1), as previously described (Ueyama et al. 2003). Briefly, the 167 bp products were digested with *Cfr10I* (Takara Shuzo Co., Kyoto, Japan) (recognition sequence = RCCGGY) and separated by 6% PAGE. The gel was stained with SybrGold (Molecular Probes, Inc., Eugene, OR, USA), and fluorescence intensity of the 137 bp band (from the first gene promoter) and of the 97 bp band (from the downstream gene promoters) was measured in the Epilight UV detector (EU-1000, Aisin Cosmos R&D, Tokyo, Japan). Gene numbers were estimated from their ratios [(137 bp + 97 bp)/137 bp]. The validity of the method was confirmed by its application to the PCR products of the cloned promoters of the first gene and downstream genes that had been mixed at various ratios (1:1–1:4). Ratio of the red/green genes was estimated from the ratio of PCR products for exon 5 (Hayashi et al. 2001). The enzyme used was the Takara *Taq* DNA polymerase (Takara). Exons 2, 3, and 4 of the red/green

Table 1 Primers used in PCR for 697 bp region and in long-range PCR

Primer	Sequence (5'-3')	Position
PF-2	GAGGAGGAGGTCTAAGTCCC	–128 to –109 (common to both promoters)
PR	GGCTATGGAAAGCCCTGTCCC	Exon 1 (+41 to +21)
DG11	TGCCCCCAAGGACCTCGTGTC	819575–819595 ^a , 856713–856733 ^a
S24	GCTTGAGCCCGAGAGTTTTAT	820450–820430 ^a , 858258–858238 ^a
PDR	CCTGAGGGTCACGGTGCTTTA	–7 to –27 (downstream gene promoter)
TX	GCTCGCCCTGTGTCTGAAATG	–31 to –11 (<i>TEX28</i> gene promoter)
FG	GAGGCGAGGCTACGGAGT	–860 to –843 (first gene promoter)
E6	GCAGTGAAAGCCTCTGTGACT	Exon 6 (3' untranslated region)
DG4	AGACAGAGTCTTGGTCTGTTG	–317 to –297 (downstream gene promoter)
S4	ATGCCTCCTGGGTTTGTCAAT	857612–857592 ^a
PF-1	CCAGCAAATCCCTCTGAGGC	–190 to –171 (common to both promoters)
PFR	CCTGAGGGTCACGGCGCTTTA	–7 to –27 (first gene promoter)
DG11-J	TGCCCGCAAGGACCTGGTGTC	<i>TEX28</i> gene in a Japanese monkey
DG11-C	TGCCCTCAAGGACCTGGTGTC	<i>TEX28</i> gene in a crab-eating monkey
S24-M	GCTTGAGCCTGAGAGTTTTAT	<i>TEX28</i> gene in the Macaque monkeys

^aFrom NT_025965 (NCBI Nucleotide Database)

genes were also amplified by the PCR as previously described (Hayashi et al. 2001; Yamade et al. 2001; Ueyama et al. 2003). PCR products were sequenced using the DYEnamic ET cycle sequencing kit (Amersham Biosciences, Amersham plc, UK) in a PRISM 310 genetic analyzer (PE Biosystems, Foster City, CA, USA).

In color-vision-deficient subjects with a single-gene array (either green or red gene only, i.e., no pTEX), the region of interest in fTEX was amplified by the PCR using primers DG11 and S24 (Table 1 and Fig. 1). The cycling parameters were: initial denaturation at 94°C for 2 min and 35 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 1.5 min, and final extension at 72°C for 2 min. After reaction, the mixture was subjected to electrophoresis on a 1% agarose gel along with the 100 bp ladder DNA size markers (Toyobo Co., Tokyo, Japan).

Fig. 1 Nucleotide sequence of the 697 bp region and its flanking regions. A part of intron 3 of *TEX28* gene (856713–858258 of NT_025965 from the NCBI Nucleotide Database) is shown. The portion between two blank lines is the 697 bp region that was absent in pTEX. Primers DG11 and S24 were used in the PCR to detect either the presence or absence of the 697 bp region. Primer S4 was used in combination with primer DG4 in the long-range PCR for the third gene of a 3-gene array. The two *Alu* repeated sequences are *doubly underlined*. The CT-rich sequence in fTEX is usually 27 bp shorter than in pTEX

TGCCCCCAAGGACCTCGTGTCCAGAAGGGAAAGAAAAACATGGAGACCTGGCATATTCAC

Primer DG11 (→)

TTCCGGGCGTTTTCAATATGACAGGAGGCTATAAGAAAATTTACGACTGCCCACCTTGTG
GATACAGCCAGAAATACAAAGCAAGAGTGAAAGTTTTGCCTACTCGGCACCTGAGTAGCT
TCTTTTTTATTCAATTTTTATTGAGGTAAATTTTCATGTAACATAAAATGAATTATTTTAA
GGTGAACAATTCAGTGGCATTTAGCACATGTGGTGTGTGCAACCACCATGGCTATTAAT
AGTTCCAAGACATTGGCCGGCTGCAGTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGA
GGCGGGCGGATCACGAGGTCAGGAGATCGAGACCATCCTGGCCAACA

CGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGTGCGCGCCT
GTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGC
TTGCAGTGAGCCGAGATCACGCCACTGCACCTCCAGCCTGGGTGACAGAACGAGACTCCGT
CTCAAAAAAAAAAAAAAAAAATTGTTCCAAAACATTTTCATCACTGCCTCGAAAAAATACTA
TTTAGCAGTCACTCCCCATTCCCCATCACCCAGTCCCTGGTAACCACCAATCTGCTTTC
TGACTTGAGTAGCTTCTTATCTCTGAACCTCTACCAGCAATTTATGAATGAGAGATGGAC
AAGACCATCAGGTATGTGGAGACGGTCCCTAGAGAAGCTGCTAGAAAAGTGACAGACAACCA
GAGTAAGGGGTAGTGGAGCACTGTAATGCTGTGGAGAGGAGAGGGCAGCAGGAATGACAA

(←) Primer S4

ACCCAGGAGGCATTGACTCAGGGAAAGAAGAGGAAGGCAGGGTCAGGTCGAGATCTTGGC
CAAATCCCTGAGCTCATCAACCTAGGGTGCTCCATTGATGAATTTATCAAGAGTCATTGG
CCGGGAGTGGTGGCTCACGCCTGTAATCACAGCACTTTGGGAGGCCAAGGCAGGTGGATC
ATTTGAGATTAGGAGTTGGAGACCAGCCTGGCCAACA

TGATGAAACCCCGTCTCTACTAAAAATATGCAAAAATTAGCTGGGCGTGGTGGTGCATGCC
TGTAATCTCAGCTACTCAGGAGGCTGAGACAGGAGAATCACTTGAATCCGGGACGCAGAG
GTTGCAGTGAGCTGAGATCAAGCCACTGCACCTCCAGCCTGGGCAGCAGAGCAAGAC
ACCACCTCAAAAAAAAAAAAAAAAAAGCCATTAACAACCTTTCTTTCTTTCTTTCTTTCTT
 --- (CT-rich sequence) --- TTTTGAGACAAGTTCTCACTCTGTCCCTCTATACCC
AGGCTGGAGCACAGTGGCACAATCATAGTTCACTGCAACATAAACTCTCGGGCTCAAGC

(←) Primer S24

When a 1,573 bp product was detected, the 697 bp region was judged to be present, and when a 876 bp product was detected, the 697 bp region was judged to be absent. These product sizes varied in length as either shorter or longer by 27 bp due to variability of the CT-rich sequence (Fig. 1), but these differences had no influence on our interpretations.

Long-range PCR

The enzyme used for the long-range PCR was the TripleMaster Polymerase (Eppendorf AG, Hamburg, Germany). The 25-μl reaction mixture contained 1× tuning buffer that was included in the kit, 500 μM each of dNTPs, 20–50 ng of template DNA, 0.4 μM each of primers, and 2 U of the polymerase enzyme. The cycling

parameters were: initial denaturation at 93°C for 3 min, ten cycles of 93°C for 15 s and 68°C (extension) for an appropriate period (at 1 min/kb), 18 cycles of 93°C for 15 s and 68°C for an appropriate period (at 1 min/kb) with an increment of 20 s/cycle, and final extension at 68°C for 10 min.

Subjects with a 2-gene array This group consisted of 94 color-normal subjects with a gene array of Red-pTEX-Green-fTEX, 36 color-vision-deficient subjects with a gene array of Green-pTEX-Green-fTEX, and 24 color-vision-deficient subjects with a gene array of Red-pTEX-Red-fTEX. The region of interest was first amplified by the PCR using primers DG11 and S24. When both 1,573 bp and 876 bp products were detected, long-range PCR was performed to examine which TEX gene had the region. Primers DG11 and PDR (Table 1) were used for amplification of pTEX (Fig. 2A), and primers DG11 and TX (Table 1) were used for amplification of fTEX

(Fig. 2A). After amplification, the mixture was subjected to electrophoresis on a 0.7% agarose gel along with the 2.5-kb ladder markers (Takara). Typical electrophoresis profiles are shown in Fig. 3 (lanes 1 and 2). The products (17,626 bp for the wild-type pTEX and 18,721 bp for the wild-type fTEX), purified from the gel using the BandPrep DNA extraction kit (Pharmacia, Uppsala, Sweden), were used as the template (usually 1/50 of the product) for the second-round PCR for the region of interest using primers DG11 and S24. The reaction condition was the same as above but for 20 cycles. The presence or absence of the 697 bp region was judged as described above.

Subjects with a 3-gene array This group consisted of 47 color-normal subjects with an array of Red-pTEX-Green-pTEX-Green-fTEX, seven color-vision-deficient subjects with an array of Green-pTEX-Green-pTEX-Green-fTEX, and six color-vision-deficient subjects with an array of Red-pTEX-Red-pTEX-Red-fTEX. First, presence or absence of the 697 bp region was analyzed, as for the 2-gene arrays. When both 1,573 bp and 876 bp products were detected, pTEX and fTEX genes were separately amplified by long-range PCR. The second-round PCR was performed using primers DG11 and S24 for either the presence or absence of the 697 bp region. When both 1,573 bp and 876 bp products were detected in pTEX, another long-range PCR for the upstream pTEX was performed using primer FG (Table 1), which is specific for the first gene promoter, and primer S24 (Fig. 2B). The reaction mixture was subjected to electrophoresis on a 0.7% agarose gel. A typical electrophoresis profile is shown in Fig. 3 (lane 3). The product (22,504 bp) was purified and used as the template in the second-round PCR using DG11 and S24. When a 1,573 bp product was detected, the upstream pTEX was judged to have the 697 bp region and, accordingly, the downstream pTEX to lack the region. When a 876 bp product was detected, the result was judged to be reverse.

Fig. 2 PCR of pigment-gene arrays. *Thin arrows* represent primers used in conventional PCRs, and *thick arrows* represent primers used in long-range PCRs. *Closed arrows* represent red genes and *open arrows* green genes. *Dotted arrows* represent TEX28 genes, and *crisscrossed arrows* represent either red or green genes. Presence of the 697 bp region is shown by *closed circles*. **A** Detection of the 697 bp region in a 2-gene array. Primers DG11 and PDR were used in the long-range PCR for pTEX, and primers DG11 and TX were used for the long-range PCR for fTEX. Using products of the long-range PCR as the template, second-round PCR was performed with primers DG11 and S24 for detection of the 697 bp region. **B** Detection of the 697 bp region in a 3-gene array. The upstream pTEX was amplified by long-range PCR using primers FG and S24. The other strategies are the same as in 2-gene arrays. **C** Analysis of the third gene in a 3-gene array. The first gene was amplified by long-range PCR using primers FG and E6. The two downstream genes were amplified together by long-range PCR using primers DG4 and E6. The third gene was amplified by long-range PCR using primers DG4 and S4. Using products of these reactions as the template, second-round PCR was performed. Although the figure shows two primers only, PF-1 and PFR, PCR for exons 2–4 was also performed, and structures of the first and third genes were definitively determined. Structure of the second gene was presumed by comparison of downstream genes and the third gene

Fig. 3 Gel electrophoresis profiles of long-range PCR products. Primers used were: DG11 and PDR for pTEX in lane 1, DG11 and TX for fTEX in lane 2, FG and S24 for the upstream pTEX in lane 3, and DG4 and S4 for the third gene in lane 4. The 2.5-kb ladders of DNA size markers are shown in lane M. Products of these long-range PCRs were subjected to electrophoresis on 0.7% agarose gels and stained with ethidium bromide

Analysis of 3-gene arrays in which gene order was unclear In color-vision-deficient subjects who had an array of Green-pTEX-Green-pTEX-Green-fTEX or Red-pTEX-Red-pTEX-Red-fTEX, the first gene was amplified by long-range PCR using primers FG and E6 (Table 1, Fig. 2C), as previously described (Oda et al. 2003). The product (15,767 bp) was used as the template in the second-round PCR for exons 2–5. Downstream genes were amplified together by long-range PCR using primers DG4 and E6. The product (13,853 bp) was used as the template in the second-round PCR for exons 2–5. The products of these second-round PCRs were subjected to sequencing. In 13 subjects, however, it could not be determined by these analyses which of the two downstream genes occupied the second position and had –71C. Gene arrays of these subjects were further analyzed in this study.

The presence or absence of the 697 bp region was first examined as described above. When the region was present in the fTEX and absent in the pTEX, the third gene of the array was amplified by long-range PCR using primer S4, which is specific for the 697 bp region (Fig. 1), and primer DG4, which is specific for the downstream gene promoter (Table 1) (Fig. 2C). The reaction mixture was subjected to electrophoresis on a 0.7% agarose gel. A typical electrophoresis profile is shown in Fig. 3 (lane 4). Though there were some thin additional bands, the main band was the expected product of 20,723 bp. The product was purified and used as the template in the second-round PCR for promoter, exon 2, exon 3, and/or exon 4. For promoter, PCR products (184 bp) using primers PF-1 and PFR (Table 1) were digested with *HhaI* (Toyobo) and analyzed by 6% PAGE. The gel was stained with SybrGold. When –71A was present, the products were expected to be cut into 169 bp and 15 bp-fragments, but when –71C was present, the 169 bp fragment was expected to be

further cut into 120 bp- and 49 bp-fragments. The primer PFR differs by one nucleotide from the primer PDR (Table 1), but we used PFR because it contains the recognition site of *HhaI* (GCGC) that is useful as an internal control of digestion. The PCR products for exons 2–4 were sequenced. A structure of the second gene was presumed from the data of the third gene.

TEX28 genes in monkeys The region of interest in the pTEX and fTEX genes was amplified by PCR using primers DG11-J and S24-M (Table 1) for the Japanese monkey and primers DG11-C and S24-M (Table 1) for the crab-eating monkey. The presence or absence of the region corresponding to the human 697 bp region was judged as for the human subjects.

Results and discussion

Differences between two *TEX28* genes

On examination of the database nucleotide sequence (NCBI Nucleotide Database, NT_025965), we found that pTEX was 731 bp shorter than the corresponding region of fTEX. Most of the difference was due to the 697 bp sequence in intron 3 present in fTEX but absent in pTEX. Figure 1 shows the nucleotide sequence of the 697 bp region with its flanking regions (the direction is opposite to that of transcription in fTEX). The 697 bp region begins within an *Alu* repeated sequence followed by a 399 bp unique sequence and ends within another *Alu* repeated sequence. Therefore, fTEX must retain the original structure, and pTEX must have undergone deletion of the 697 bp region, as depicted in Fig. 4 (processes 1–3).

Frequency of the 697 bp region

We investigated whether all of fTEX and none of pTEX genes have the 697 bp region in 94 color-normal Japanese males with a 2-gene array, Red-pTEX-Green-fTEX. Figure 5A shows the three haplotypes that were detected. In the pTEX genes, 43% (40/94) had the 697 bp region. Alternatively, in the fTEX genes, 97% (91/94) had the region. The other haplotype in which the 697 bp region was present in pTEX and absent in fTEX was not found in the 94 subjects.

Three-gene array in color-normal subjects

We examined a 3-gene array, Red-pTEX-Green-pTEX-Green-fTEX, in 47 color-normal male Japanese subjects. The five haplotypes that were detected are shown in Fig. 5B. In the pTEX genes, 10% (9/94) had the 697 bp region. In the fTEX genes, 87% (41/47) had the 697 bp region.

Nathans et al. (1986a) suggested that a normal 3-gene array would have arisen through unequal homologous

Fig. 4 Deletion of the 697 bp region in the pTEX gene during evolution of red/green visual pigment genes. The ancestral pigment gene (crisscrossed arrow without showing each of six exons) was followed by the *TEX28* gene (fTEX, dotted boxes and an arrowhead representing exons) having the 697 bp region (closed circles) in intron 3. At process 1, a segment between vertical bars, which contained the whole of the pigment gene and exons 2–5 of *TEX28* gene, was duplicated generating another pigment gene and a nonfunctional *TEX28* gene (pTEX) between the pigment genes. After process 2, in which these two pigment genes evolved into red (closed arrow) and green (open arrow) genes, some of the pTEX gene underwent a deletion of the 697 bp region (process 3). At process 4, another gene duplication may have occurred between two vertical broken bars, though their positions are uncertain, generating another green gene and another pTEX gene without the 697 bp region

crossover between normal 2-gene arrays. Figure 6A shows an example of such crossover, which yields a normal array with two green genes (“a–b” array) and an array with a single red gene (“b–a” array). If this pathway is correct, the frequency of the 697 bp region in the upstream pTEX in normal 3-gene arrays should be approximately 43%, which we estimated from the results in normal 2-gene arrays (Fig. 5A). However the observed frequency was 11% (5/47, Fig. 5B). Moreover, an example of such crossover in Fig. 5A shows that all the downstream pTEX in the “a–b” array have the 697 bp region irrespective of crossover points. Its probability is estimated to be 42% (0.97×0.43), which is already much higher than the observed frequency of 8.5% (4/47, Fig. 5B).

Normal arrays with two green genes have a frequency of 43% (Drummond-Borg et al. 1989), which is out of balance with the, at most, 1.97% frequency of single-red-gene arrays, as estimated from the frequency of dichromacy in the deutan color-vision deficiency (Drummond-Borg et al. 1989). The discrepancy may indicate that either single-red-gene arrays have been subjected to negative selection during human evolution,

or the pathway shown in Fig. 6A may only partially contribute to generation of normal 3-gene arrays.

Our results suggest a new pathway where normal arrays with multiple green genes may have arisen through second-round gene duplication, which occurred predominantly in the array where the pTEX had already undergone deletion of the 697 bp region, as depicted in Fig. 4 (process 4). We have no explanations for such a limited duplication, but a second-round gene duplication itself has been suggested from gene analysis of color-vision-deficient subjects. Subject A67, for example, had a 4-gene array, and all three downstream genes had the same 11 bp deletion in exon 3 (Ueyama et al. 2004). Subject A164 had a 3-gene array, and both downstream genes had the same point mutation in exon 6 (Arg330Gln, CGA → CAA, Ueyama et al. 2002).

Onishi et al. (2002) reported that 93% (124/133) of the crab-eating monkeys they studied had a 2-gene array: Red–Green. The array of a single red gene followed by multiple green genes was restricted to the six monkeys in the southern part of Thailand. The frequency is quite different from that in humans, which is only 19% (25/130, Drummond-Borg et al. 1989). We investigated whether a crab-eating monkey had the 697 bp region, together with a Japanese monkey. Each monkey seemed to have a pair of 2-gene arrays and had 697 bp (+) *TEX28* genes only. These results indicated that both pTEX genes in each monkey had the 697 bp region. It may be of interest to study pTEX genes in the six crab-eating monkeys with multiple green genes.

Single-gene array in color-vision-deficient subjects

We next examined color-vision-deficient subjects with a single-gene array to determine whether their fTEX genes have the 697 bp region. A single-gene array is thought to result from unequal homologous crossover between two

Fig. 5 Summary of the insertion/deletion polymorphism in arrays of color-normal and color-vision-deficient subjects. The number of subjects is shown beside each array. *Closed arrows* indicate red genes and *open arrows* indicate green genes. *Dotted arrows* indicate both pTEX and fTEX genes. *Closed circles* represent the 697 bp region

pigment gene arrays. When it occurs upstream of exon 5, one of the arrays becomes a single-green-gene array (Fig. 6B). The frequency of the 697 bp region in fTEX of the subjects with a single-green-gene array was 83% (25/30) (Fig. 4C), which was not so different from the expected frequency of 97% ($0.54 + 0.43$, Fig. 5A).

When unequal homologous crossover occurs downstream of exon 5, one of the arrays becomes a single-red-gene array (Fig. 6A). The frequency of the 697 bp region in fTEX of the subjects with a single-red-gene array was 66% (74/112) (Fig. 5C). The distance from the end of exon 5 to the beginning of the 697 bp region (8,881 bp) is about a half of that from the end of the 697 bp region

to the beginning of pTEX (17,030 bp). Therefore, in the crossover that generates a single-red-gene array, one third is presumed to occur upstream of the 697 bp region and two thirds to occur downstream of the region. The observed frequency of 66% is consistent with the presumed frequency of 61% [$(0.54 \times 1/3) + 0.43$, Fig. 5A].

Two-gene array in color-vision-deficient subjects

In the color-vision-deficient subjects with a 2-green-gene array, 44% (16/36) of the pTEX and 97% (35/36) of the fTEX had the 697 bp region (Fig. 5D). The frequency of

Fig. 6 Unequal homologous crossover between pigment gene arrays. *Closed arrows* represent red genes, and *open arrows* indicate green genes. *Dotted arrows* indicate both pTEX and fTEX genes. *Closed circles* represent the 697 bp region. "X" represents positions of crossover. **A** Unequal crossover downstream of exon 5 between two normal arrays, a and b, yields a normal array with two green genes (a-b) and an array of single red gene (b-a). **B** Unequal crossover upstream of exon 5 between two normal arrays yields two defective arrays, one of which is a single-green-gene array. **C** Unequal crossover upstream of exon 5 between a normal 2-gene array and a normal 3-gene array yields two defective arrays, one of which is a 2-green-gene array. **D** Unequal crossover downstream of exon 5 between a single-green-gene array, and a normal 2-gene array yields a 2-green-gene array and a single-red-gene array. **E** Unequal crossover upstream of exon 5 between a single-red-gene array and a normal 2-gene array yields a 2-red-gene array and a single-green-gene array. **F** Unequal crossover downstream of exon 5 between a 3-green-gene array of Red-Red-Green and a normal 2-gene array yields a 2-red-gene array and a normal 3-gene array

44% is much higher than expected. If unequal homologous crossover between normal 2-gene and normal 3-gene arrays is the pathway for generation of a 2-green-gene array, which is depicted in Fig. 6C, the frequency of the 697 bp region in pTEX is estimated from the results (Fig. 5B) to be approximately 8.5%. We propose that unequal homologous crossover between a single-green-gene array and a normal 2-gene array may be a major pathway for a 2-green-gene array (Fig. 6D).

Unequal homologous crossovers between a single-red-gene array and a normal 2-gene array (Fig. 6E) and between a 3-gene array of Red-Red-Green and a normal 2-gene array (Fig. 6F) are possible pathways for

generation of a 2-red-gene array. The frequency of the 697 bp region in pTEX is expected to be 43% in the former (Fig. 5A) but to be 11% in the latter (5/47, Fig. 5B). The observed frequency of 75% (18/24, Fig. 5D) suggests that the former is the major pathway for generation of a 2-red-gene array.

Presumption of pigment gene structure using the 697 bp region

There is good evidence for the hypothesis that only the first two genes of the pigment gene array are expressed (Hayashi et al. 1999). Therefore, information about the first two genes of an array is necessary to predict color-vision phenotypes. Up until now, gene analysis has been difficult in arrays consisting of more than two genes.

In the red/green visual pigments, mainly the exon-5-encoded sequence contributes to the difference in $\lambda_{\max}$, but exon-2-encoded, exon-3-encoded, and exon-4-encoded sequences cause minor spectral shifts. Therefore, if none of these three exons differ between the first and second genes of the array, the phenotype is predicted to result in dichromacy, which is the most severe form of color-vision deficiency. If any of these three exons differs, the phenotype is predicted to be anomalous trichromacy, which is a less severe form.

The most downstream gene of the array has been analyzed using a long-range PCR method (Hayashi et al. 1999; Oda et al. 2003), but it was for the exon 5 only, because one of the primers had to be designed near exon 1 of fTEX gene. The distance from exon 5 of the pigment gene to the exon 1 of fTEX gene was near the 28,000 bp limit of amplification of long-range PCR. In all subjects listed in Fig. 7, none of the pTEX and all of the fTEX had the 697 bp region. Using the 697 bp region-specific primer S4, we could amplify the whole of the third gene of a 3-gene array, including its promoter, and predict a structure of the second gene.

We reported that an A-71C substitution in the promoter of the second gene of the array was associated with color-vision deficiency (Ueyama et al. 2003). All subjects in Fig. 7 had one downstream gene with -71A and the other with -71C. When the second gene of the array has -71C, its expression may be repressed causing dichromacy, irrespective of differences in exons 2-4 between the first and second genes.

Results are summarized in Fig. 7. In the first seven subjects of Fig. 7, it was unclear whether or not the first and second genes had the same structure. Our present study indicated that in A158, exon 3 was different between the first and second genes, but the second gene probably had -71C (Fig. 8, lane 1), which may account for the subject's dichromatic phenotype. In the other six subjects, A32, A291, A321, 329, A380, and A419, we found no differences between the first and second genes, consistent with their dichromatic phenotype. The nucleotide at -71 in the third gene was C in these six subjects (Fig. 8, lanes 2-7).

Fig. 7 Summary of color-vision-deficient subjects with a 3-gene array in which gene order was unclear. For the pigment genes, *boxes and arrowheads* represent exons. Exons 1 and 6 are all crisscrossed because there are no differences between red and green genes. *Closed boxes* for exon 2 mean that codon 65 was for Thr and codon 111 was for Ser. *Open boxes* for exon 2 mean that codon 65 was for Ile and codon 111 was for Tyr. For exon 3, *closed boxes* mean that codon 180 was for Ser and *open boxes* mean that it was for Ala. *Closed boxes* for exon 4 mean that codon 230 was for Ile and codon 233 was for Ala. *Open boxes* for exon 4 mean that codon 230 was for Thr and codon 233 was for Ser. *Closed boxes* for exon 5 mean that codon 277 was for Tyr and codon 285 was for Thr. *Open boxes* for exon 5 mean that codon 277 was for Phe and codon 285 was for Ala. *Dotted arrows* represent *TEX28* genes, and *closed circles* indicate the 697 bp region

Fig. 8 Analysis of the third-gene promoter. The third gene of a 3-gene array was amplified by long-range PCR using primers DG4 and S4, and the product was used for the second-round PCR for promoter using primers PF-1 and PFR. The product (184 bp) was digested with *HhaI* and analyzed by PAGE. The gel was stained with SybrGold. M, 100 bp ladders of DNA size markers; U, PCR product before *HhaI*-digestion; C, control PCR products after *HhaI*-digestion, one with -71A yielding a 169 bp band and the other with -71C yielding 120 bp and 49 bp bands. Subjects in lanes: 1, A158; 2, A32; 3, A291; 4, A321; 5, A329; 6, A380; 7, A419; 8, A8; 9, A19; 10, A84; 11, A427; 12, A316; 13, A384

In the remaining six subjects, A8, A19, A84, A427, A316, and A384, the third genes had -71C (Fig. 8, lanes 8–13) and accordingly, the second genes were predicted to have -71A. The phenotype of these subjects predicted from gene analysis was consistent with their actual anomalous trichromacy phenotype. Since the 697 bp

region is mostly present in fTEX and mostly absent in pTEX in arrays consisting of more than two genes, our method using the insertion/deletion polymorphism will be useful for analysis of 3-gene arrays.

Acknowledgements We express our sincere gratitude to Professor Iwao Ohkubo at Shiga University of Medical Science (Department of Medical Biochemistry) for his helpful advice and discussions. We thank Mr. Masashi Suzuki, Central Research Laboratory at Shiga University of Medical Science, for his technical help (sequencing). This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (14571667 to S.Y. and 12770953 to S.O.).

References

- Asenjo AB, Rim J, Oprian DD (1994) Molecular determinants of human red/green color discrimination. *Neuron* 12:1131–1138
- Deeb SS, Lindsey DT, Hibiya Y, Sanocki E, Winderickx J, Teller DY, Motulsky AG (1992) Genotype-phenotype relationships in human red/green color-vision defects: molecular and psychophysical studies. *Am J Hum Genet* 51:687–700
- Drummond-Borg M, Deeb SS, Motulsky AG (1989) Molecular patterns of X chromosome-linked color vision genes among 134 men of European ancestry. *Proc Natl Acad Sci USA* 86:983–987
- Hanna MC, Platts JT, Kirkness EF (1997) Identification of a gene within the tandem array of red and green color pigment genes. *Genomics* 43:384–386
- Hayashi T, Motulsky AG, Deeb SS (1999) Position of a 'green-red' hybrid gene in the visual pigment array determines colour-vision phenotype. *Nat Genet* 22:90–93
- Hayashi S, Ueyama H, Tanabe S, Yamada S, Kani K (2001) Number and variations of the red and green visual pigment genes in Japanese men with normal color vision. *Jpn J Ophthalmol* 45:60–67
- Merbs SL, Nathans J (1992) Absorption spectra of the hybrid pigments responsible for anomalous color vision. *Science* 258:464–466
- Nathans J, Thomas D, Hogness DS (1986a) Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* 232:193–202
- Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS (1986b) Molecular genetics of inherited variation in human color vision. *Science* 232:203–210

- Nei M, Zhang J, Yokoyama S (1997) Color vision of ancestral organisms of higher primates. *Mol Biol Evol* 14:611–618
- Oda S, Ueyama H, Nishida Y, Tanabe S, Yamade S (2003) Analysis of L-cone/M-cone visual pigment gene arrays in females by long-range PCR. *Vision Res* 43:489–495
- Onishi A, Koike S, Ida-Hosonuma M, Imai H, Shichida Y, Takenaka O, Hanazawa A, Komatsu H, Mikami A, Goto S, Suryobroto B, Farajallah A, Varavudhi P, Eakavhibata C, Kitahara K, Yamamori T (2002) Variations in long- and middle-wavelength-sensitive opsin gene loci in crab-eating monkeys. *Vision Res* 42:281–292
- Ueyama H, Kuwayama S, Imai H, Tanabe S, Oda S, Nishida Y, Wada A, Shichida Y, Yamade S (2002) Novel missense mutations in red/green opsin genes in congenital color-vision deficiencies. *Biochem Biophys Res Commun* 294:205–209
- Ueyama H, Li Y-H, Fu G-L, Lertrit P, Atchaneeyasakul L, Oda S, Tanabe S, Nishida Y, Yamade S, Ohkubo I (2003) An A–71C substitution in a green gene at the second position in the red/green visual-pigment gene array is associated with deutan color-vision deficiency. *Proc Natl Acad Sci USA* 100:3357–3362
- Ueyama H, Kuwayama S, Imai H, Oda S, Nishida Y, Tanabe S, Shichida Y, Yamade S (2004) Analysis of L-cone/M-cone visual pigment gene arrays in Japanese males with protan color-vision deficiency. *Vision Res* 44:2241–2252
- Yamade S, Hayashi S, Ueyama H, Tanabe S, Hukami K, Ichikawa K, Tachibana M (2001) Red–green pigment gene analysis as a clinical diagnostic tool. *Color Res Appl* 26:S89–S92