DNA Secondary Structures and the Evolution of Hypervariable Tandem Arrays*

M. Neale Weitzmann‡, Kerry J. Woodford‡, and Karen Usdin§

From the Section on Genomic Structure and Function, Laboratory of Molecular and Cellular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0830

Tandem repeats are ubiquitous in nature and constitute a major source of genetic variability in populations. This variability is associated with a number of genetic disorders in humans including triplet expansion diseases such as Fragile X syndrome and Huntington's disease. The mechanism responsible for the variability/instability of these tandem arrays remains contentious. We show here that formation of secondary structures, in particular intrastrand tetraplexes, is an intrinsic property of some of the more unstable arrays. Tetraplexes block DNA polymerase progression and may promote instability of tandem arrays by increasing the likelihood of reiterative strand slippage. In the course of doing this work we have shown that some of these tetraplexes involve unusual base interactions. These interactions not only generate tetraplexes with novel properties but also lead us to conclude that the number of sequences that can form stable tetraplexes might be much larger than previously thought.

Tandemly repeated DNA sequences are distributed widely in nature and may constitute as much as 10% of the human genome (1). They are sometimes referred to as satellites, mini-satellites, or microsatellites, depending on their repeat size or genome (1). They are sometimes referred to as satellites, mini-

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ These authors have contributed equally to this work.

§ To whom correspondence should be addressed: Bldg. 8, Rm. 202, National Institutes of Health, 8 Center Dr. MSC 0830, Bethesda, MD 20892-0830. Tel.: 301-496-2189; Fax: 301-402-0240; E-mail: ku@helix.nih.gov

1 The abbreviations used are: HVRs, hypervariable repeats; PCR, polymerase chain reaction; DMS, dimethyl sulfate; dd, dideoxy.

show polarity. Tandem arrays are also prone to the acquisition of point mutations, and the distribution of these mutations shows a similar polarity (9, 12, 16, 17). This has led to the suggestion that either flanking sequences are important in imparting polarity to an otherwise non-polar process (12) or a mechanism that has an inherent polarity such as replication slippage (16) is involved. However, many of the most hypervariable arrays show a many-fold increase in repeat number that is thought to take place within the space of only a few cell divisions (18). Such a large increase in repeat number cannot be accomplished by a single strand slippage or recombinational event, and it has been suggested that in such cases some specialized mutational mechanism must be active (19, 20).

Many hypervariable sequences that have been described are G + C-rich and show a strand asymmetry in that one strand is predominantly G-rich and the other C-rich (21). It had been suggested that these sequences contained a \( \chi \)-like sequence that could account for the observed variability by promoting recombination (10). However, many of the more recently identified hypervariable sequences lack a discernible \( \chi \)-like motif. We had previously found that a hypervariable sequence, the CGG repeat in the human \( FMR1 \) gene that undergoes triplet expansion to result in Fragile X syndrome (22, 23), forms a series of intrastrand tetraplexes at physiological temperatures, pH, and ionic strengths (24). This occurs despite the fact that this sequence was one-third Cs, and this C-richness would be expected to reduce tetraplex stability. We have now tested a series of other highly hypervariable tandem repeats (Table I) for the ability to form intrastrand tetraplexes using a \( K^+ \)-dependent arrest of DNA synthesis assay that we have recently developed (25). These sequences are also G + C-rich hut, like the CGG-repeat at the \( FMR1 \) locus, contain a number of non-G bases. We have found that the ability to form intrastrand tetraplexes is a shared property of all of these sequences. This, together with the observation that other hypervariable tandem arrays form hairpins (24, 26–32), or triplexes (33), supports the idea that DNA secondary structure may play a major role in the generation and evolution of tandem arrays.

MATERIALS AND METHODS

Clone Construction—Oligonucleotides containing hypervariable repeat units were synthesized on an ABI 381A oligonucleotide synthesizer using standard phosphoramidite chemistry and cloned into the plasmid pMS189 as described previously (24, 34). Plasmids were replicated in \textit{Escherichia coli} MBM7070, isolated by alkaline lysis, and purified by CsCl gradient centrifugation according to standard procedures.

Intrastrand Tetraplex Assay—Hypervariable sequences were tested for the ability to block DNA synthesis reactions as follows (25). Sequencing primer was phosphorylated with \( \gamma^\text{32P} \)ATP (DuPont NEN, 3000–6000 Ci/mmol) using \( T\lambda \) polynucleotide kinase (Epitcore Technologies, Inc.), and a buffer containing 50 mM Tris-HCl, pH 8.0, and 10 mM MgCl\(_2\). Reaction mixtures (total volume 6 \( \mu \)l) contained 0.2–2 \( \mu \)M template, 0.16 mM of the primer SupFUr (5'-ATGCTTTTACTGGCCCT-GCT-3'), 10 \( \mu \)M dNTPs, one of the following dideoxynucleotides at the
concentration indicated in parentheses: ddATP (0.3 mM), ddGTP (0.017 mM), ddCTP (0.2 mM), ddTTP (0.6 mM), 50 mM Tris-HCl, pH 9.5, 2.5 mM MgCl₂, 5 units of Taq polymerase (Life Technologies, Inc.), and where indicated 50 mM monovalent cation. Reaction mixtures were subjected to 30 rounds of heating and cooling in a Perkin-Elmer PCR machine for 30 s at 95 °C, 30 s at 55 °C, and 30 s at 72 °C. The reaction was terminated by the addition of one-half volume of stop buffer containing 50% (v/v) formamide, 10 mM EDTA, pH 9.5, 10 mM NaOH, 0.1% xylene cyanol, and 0.1% bromphenol blue, and the mixtures were heated at 90 °C for 5 min prior to electrophoresis on a 6.5% polyacrylamide sequencing gel. The sequence located between the sequencing primer and the repeat on the template strand is 5′-CTCAGGTCAAGTGTTACCTTTACACGGCGGCGTCTATTTGATGTGCGGCCCAGCTTTCCCGATAAGGGG-3′.

**Preparation of 7-Deazaadenosine templates by PCR**—Templates containing guanine or 7-deazaadenosine were prepared by PCR amplification of plasmids containing the HVR of interest using the primers AMP2 (9′-GGCCACGGAAAGTTGGA-3′) and supFR1 (5′-GATCGAATTCGTCGACGATGTTGGGGGAA-3′) which flank the HVR. The primer binding sites are located about 500 bases apart, the precise distance depending on the template, with the repeat being located about halfway between the two primer binding sites. Reaction mixtures containing 10 ng of plasmid template DNA containing the repeat of interest: 1 μM each of AMP2 and supFR1; 2.5–5 units of Taq polymerase (Life Technologies, Inc.); 50 mM Tris-HCl, pH 8.0; 10 mM MgCl₂; 100 or 160 μM each of dATP, dGTP, dCTP, and either ddGTP or 7-deaza-dGTP were prepared. They were then overlaid with a drop of mineral oil and subjected to 30 rounds of heating and cooling in a Perkin-Elmer PCR machine for 30 s at 95 °C, 30 s at 55 °C, and 30 s at 72 °C. The PCR products were purified on a 5% polyacrylamide gel and used as templates in the tetraplex assay described above.

**DMS Protection Assays**—Dimethyl sulfate (DMS) protection assays were performed on gel-purified oligonucleotides using the method of Williamson et al. (35) with slight modifications. End-labeled oligonucleotide (1–5 ng per reaction) was resuspended in 18 μl of TE buffer and heated to 95 °C for 2 min. Following this, the samples were precipitated twice with 1.2 ml of butanol. The samples were dried under vacuum, redissolved in 20 μl of 95% (v/v) formamide, 5 mM EDTA, pH 9.5, 5 mM NaOH, 0.05% xylene cyanol, 0.05% bromphenol blue, denatured for 5 min at 90 °C, and run on a 20% sequencing gel. Gels were covered with plastic wrap and exposed to x-ray film overnight at −20 °C.

### RESULTS

Intrastrand tetraplexes form when four G-rich motifs on a single strand interact to form a series of tetrad (36–39). A series of stacked tetrad creates a hollow stem or cylinder. This stem is bounded by three loops formed by bases between the series of stacked tetrads creates a hollow stem or cylinder. This single strand interact to form a series of tetrads (36–39). A template containing five G₄ tetrads (shown as gray parallelograms) forms a tetraplex cavity (not to scale).

**FIG. 1.** The K⁺-dependent block to DNA synthesis assay for tetraplex formation. Diagrammatic representation of the tetraplex arrest assay on a template containing a generic intrastrand tetraplex containing five G₄ tetrads (shown as gray parallelograms). The loops L₁, L₂, and L₃ each contain three unspecified bases (N). DNA synthesis starts 3′ of the tetraplex-forming region and proceeds in a 5′ to 3′ direction toward the tetraplex. The front end of the polymerase is represented by the diagonally striped bullet and the nascent DNA strand by the dashed line. The site of premature chain termination that would result from the formation of the tetraplex on the template strand is indicated by the filled arrow. Inset, a G₄ tetrad with a K⁺ ion situated within the tetrad cavity.

One of the most unstable loci thus far identified in any organism is the mouse minisatellite locus *Mso-hm*, which has a germline mutation rate of 2.5% per gamete and which shows frequent intergenerational changes of a kilobase or more (44). This locus contains from 200 to >1000 repeats of the pentamer 5′-CAGGG-3′. A template containing eight CAGGG repeats was tested for the ability to form a K⁺-dependent block to DNA synthesis. Two distinct non-dideoxynucleotide-mediated chain termination products are seen at the 3′ end of the repeat tract in the presence of 50 mM KCl when the G-rich strand is used as a template (Fig. 2). The more prominent of the two products (filled arrow) corresponds to a block to DNA synthesis just 3′ of the first G residue in the primer 5′-CAGGG-3′ repeat on the template. The second product (open arrow) corresponds to premature chain termination one base 3′ of this one. A series of weaker stops are seen at corresponding positions in the next four repeats. A smaller amount of premature termination is also observed in the presence of 50 mM NaCl, but none is observed in the absence of cation or in the presence of LiCl.
DNA Structures and G + C-rich Hypervariable Regions

FIG. 2. Tetraplex assay of the Ms6-hm locus. The assay was carried out on the G-rich and the C-rich strands, in the presence of the indicated monovalent cations as described under “Materials and Methods.” The lane markers T, C, G, and A indicate the bases on the template strand. The bracket on the left side of the figure indicates the tandem repeat, and the open and filled arrows mark the positions of sites of the major monovalent cation-dependent sites of premature chain termination.

RbCl, CsCl, or NH₄Cl. Since metal binding sites on a hairpin are equally accessible to all cations, and the affinity of cations for binding sites on DNA decreases slightly with increasing metal ion radius (45), the cation specificity is inconsistent with the blocks being due to hairpin formation. No block to DNA synthesis is seen when the complementary strand is used as a template (Fig. 2, right panel) or when single-stranded phage DNA is used as a template (data not shown), ruling out structure triplexes that involve interactions between the template and its complementary strand (46). Arrest of DNA synthesis is seen when these repeats are cloned into other vectors (data not shown), indicating that flanking sequences are not involved. Blockage is also independent of template concentration over a wide range (data not shown) indicating that the blocks do not involve interactions between two or more template strands but are due to the formation of intrastrand structures.

The properties of both the Na⁺- and the K⁺-dependent DNA synthesis arrest sites including the position of the blocks to DNA synthesis, the template concentration independence, and the strand specificity, are most consistent with intrastrand tetraplex formation. The major stop reflects the most stable tetraplex(es) involving the maximum number of repeats. The less prominent stops at subsequent repeats reflect a series of tetraplexes that presumably involve a smaller number of repeats. In addition to these monovalent cation-dependent stops, a smaller amount of cation-independent premature chain termination is seen at the second G of every repeat. These stops are even more marked in both guanine and 7-deazaguanine containing linear templates (Fig. 3), and this is paralleled by a hypersensitivity of that G to methylation by DMS (see Fig. 4). We hypothesize that these phenomena may be related to a conformational peculiarity of the DNA backbone of this region.

To confirm that polymerase arrest in the presence of K⁺ and Na⁺ is related to tetraplex formation, the polymerase chain reaction (PCR) was used to generate templates containing either guanine or 7-deazaguanine. These templates were then tested for the ability to cause K⁺/Na⁺-dependent DNA synthesis arrest. Since 7-deazaguanine cannot act as an N7 donor needed to form G tetrads, substitution of all guanine residues with 7-deazaguanine should abolish the K⁺/Na⁺-dependent polymerase blocks. As can be seen in Fig. 3, this is precisely what happens. The PCR template in which all the Gs have been replaced by 7-deazaguanine have lost all the K⁺/Na⁺-dependent blocks to DNA synthesis, whereas the PCR template containing guanines produced the same blocks to DNA synthesis seen on the circular templates (Fig. 3).

DMS treatment of an oligonucleotide containing the HVR was also carried out. Since Gs involved in tetrads do not have their N7 positions exposed, they are protected from modification by DMS. In theory, Gs in tetrads are completely protected from DMS, whereas Gs in the loops of the tetraplex that are not involved in intraloop or interloop interactions should be DMS-reactive (24, 48). In practice, the picture is not always so clear, and this represents a very real limitation on the value of this technique. For example, if a tetraplex is not very stable and is formed by only a small fraction of the molecules in the population, this may produce a pattern of DMS modification in which only partial protection of Gs is apparent. In addition,
many tetraplex-forming sequences show conformational complexity that can complicate DMS data interpretation, since a base protected in one structure may be exposed in another. Since the fraction of molecules in the population that form a K\(^+\)-dependent block to DNA synthesis in the case of the mouse Ms6-hm HVR is small, we would expect to see some DMS protection, but this protection would not be complete. This is in fact the case (Fig. 4). After normalizing the K\(^+\) and K\(^-\)-free reactions to a G outside of the HVR (indicated by an asterisk in Fig. 4) we can see that Gs within the HVR show less DMS reactivity when K\(^+\) is present than when it is absent. While not definitive, these data are consistent with our other data and support the idea that the mouse Ms6-hm HVR is capable of tetraplex formation.

Why a Na\(^+\)-induced polymerase block is seen only with this sequence and not other tetraplexes we have tested (24, 25, 34, 47) is not clear, but preliminary evidence suggests that it is related to the involvement of adenines in the structure since the sequence (CTGGG\(\)\textsubscript{12}) shows K\(^+\)-dependent but not Na\(^+\)-dependent DNA polymerase arrest (data not shown). However, the mere presence of adenines is not sufficient to elicit a Na\(^+\) stop since not all A containing templates show such stops (Fig. 5). Rather we believe the Na\(^+\) effect is related to a specific hydrogen bonding interaction in which As are involved. The molecular basis of the Na\(^+\) effect is currently under investigation.

Tandem arrays of the repeat 5'-TGG-3' are polymorphic (49), as are a mixture of the triplets AGG and TGG (50). As with the mouse Ms6-hm minisatellite, we found that a template containing (TGG)\textsubscript{20} blocked DNA synthesis in a K\(^+\)-dependent manner (Fig. 5A), producing a series of premature chain termination products corresponding to arrest opposite the T residues of repeats 13–20 in the (TGG)\textsubscript{20} tract. No blocks are seen when the complementary pyrimidine-rich strand was used as template (Fig. 5A). The blocks to DNA synthesis disappear when 7-deazaguanine is incorporated into the template strand (Fig. 5A). A single novel weak stop (open circle) is observed at the second guanine base in repeat 20 on PCR templates containing 7-deazaguanine. This stop is also seen in PCR templates containing guanines and is not dependent on monovalent cation since it is seen in the absence of KCl (data not shown). Since this stop is unique to the PCR templates, it is not affected by substitution of Gs by 7-deazaguanine, and it is not related to the presence of K\(^+\), we presume that it reflects some aspect of the linear templates that is not related to tetraplex formation. Most of the guanines in the TGG repeat are also either fully or partially protected from methylation by DMS (Fig. 6, left panel), consistent with tetraplex formation.

We have previously shown that a (CGG)\textsubscript{16} tract blocks DNA synthesis in a similar manner producing eight premature chain termination products opposite C residues at the 3' end of the CGG tract (24). The similarity in both the pattern of polymerase arrest and DMS protection leads us to think that the tetraplexes formed by these sequences could be very similar. Such tetraplexes may contain G\textsubscript{4} tetrads interspersed with pyrimidines or a smaller number of G\textsubscript{4} tetrads interspersed with a mixture of Gs and either T or C. We have previously shown that an AGG triplet does not destabilize a CGG-containing tetraplex (24). It is therefore reasonable to assume then that a mixture of AGGs and TGGs would also form a tetraplex.

We also tested repeats with the sequence 5'-GGGAGGG-3'. Between 1 and 22 repeats of this unit are found upstream of the Huntington’s disease gene in humans (51). A template containing 2.5 repeats of this sequence produces a complex pattern of premature chain terminations. There is at least one strong strand-specific K\(^+\)-dependent block to DNA synthesis and a number of other more minor ones. A small amount of monovalent cation-independent polymerase arrest is seen at the 3' end of the D4S43 tract. This may be due to formation of a small amount of tetraplex in the absence of monovalent cation or the formation of another structure such as a hairpin that forms independently of added monovalent cation. A significant amount of monovalent cation-independent arrest is seen in the middle of this tract (indicated by the dashed line in Fig. 5B). This block is consistent with triplex formation between the G-rich template and the nascent strand (52). Any or all of these blocks to DNA synthesis could explain the difficulties reported in amplifying this region by PCR and the observation that incorporation of 7-deazagTTP is able to correct this problem (51). Once again, the K\(^+\)-dependent blocks disappear when other monovalent cations are substituted for K\(^+\), or when K\(^+\) is omitted, and no K\(^-\)-dependent stops are seen when the complementary pyrimidine-rich strand is used as a template.

Substitution of guanines in the template with 7-deazaguanine eliminates the K\(^+\)-dependent blocks to DNA synthesis (Fig. 5B). The K\(^+\)-independent polymerase arrest observed midway through the sequence is also eliminated, supporting the hypothesis that this stop may represent a purine:purine:pyrimidine triplex formed between the template and the nascent strand produced in the assay. This HVR shows a pattern of DMS modification with alternating regions of DMS protection and DMS reactivity in the presence of K\(^+\) (Fig. 6). This contrasts with the almost uniform reactivity of Gs in the absence of K\(^+\). Some of the most protected bases show a DMS reactivity indistinguishable from background. Both the 7-dea-
zaguanine substitution data and the DMS protection data are thus consistent with tetraplex formation.

Four repeats from the type I diabetes-linked hypervariable region in the human insulin promoter also produce a number of $K^+\text{-dependent}$ blocks to DNA synthesis consistent with an array of different tetraplexes (Fig. 5C). These blocks are eliminated by substitution of guanine with 7-deazaguanine and are not observed on the complementary pyrimidine-rich strand. A number of Gs in the HVR are as reactive with DMS as a reference base outside the repeat (indicated with an asterisk in Fig. 6, right panel). These Gs are separated by regions of protected Gs in which no reactivity can be seen above background. Based on indirect evidence from gel electrophoretic mobility assays, and using enzymatic and chemical probes, it had been suggested that this region is able to form a series of intramolecular tetraplexes (43, 53, 54). Our data support this claim.

**DISCUSSION**

Our observations suggest that the ability to form an intrastand tetraplex in vitro is a common feature of a number of hypervariable sequences including the mouse minisatellite at the Ms6-hm locus which is one of the most hypervariable sequences thus far described (44). The tetraplex formed by the repeats in the Ms6-hm tandem array is unusual in that it can be stabilized by $Na^+$ as well as $K^+$, albeit with lower efficacy. This contrasts with our observations that all other tetraplexes that we have tested are seen only in the presence of $K^+$ (24, 25, 34, 40, 47). Since the ionic radius of $Na^+$ is smaller than that of $K^+$, it may be that the Ms6-hm tetraplex has smaller internal dimensions than the other previously described tetraplexes.

This interpretation is consistent with the fact that other monovalent cations such as $Rb^+$, $Cs^+$, and $NH_4^+$ do not result in a block to DNA synthesis in our assay, since these ions have radii that are all larger than that of $K^+$. $Li^+$, on the other hand, is much smaller than $Na^+$ and may still be too small to form the coordination complex that is important in stabilizing these types of structures (41). Our assay might thus be useful in distinguishing between different kinds of tetraplexes such as those that are $K^+$-specific and that correspond to previously described G$_4$ tetrad containing tetraplexes and those that are also seen in the presence of other cations, specifically $Na^+$, that may represent a novel class of tetraplex with different base interactions and thus different properties.

Since we have shown previously that the amount of $K^+$ used

![Fig. 5. Tetraplex formation by different HVRs.](image link)

![Fig. 6. Dimethyl sulfate protection assays of the (TGG)$_n$, the D4S43 HVR, and the insulin HVR.](image link)
in this assay represents saturating amounts of cation for tetraplex formation (24), it is likely therefore that the same pattern of polymerase pausing/tetraplex formation would be seen at physiological [K+] which typically is around 150 mM in mammalian cells (55). Tetraplex formation in vivo would require these regions to be transiently unpaired at some time. This might occur during DNA replication or on extrusion from otherwise duplex molecules (53, 56) any time during the cell cycle. In eukaryotic cells it is thought that only relatively small regions of DNA are unpaired during replication, although it has been suggested that many hundreds of bases can be unpaired under certain circumstances (57). Direct evidence for an altered structure in vivo has been obtained for one of these sequences, that of the human insulin HVR (58), suggesting that formation of DNA tetraplexes by the hypervariable sequences described here might in fact be possible. The fact that a variety of tetraplex-binding proteins have been isolated from eukaryote cells (59–65) supports the idea that tetraplexes can form in vivo. The HVRs we have tested are much shorter than those actually found at their specified loci on chromosomes. Therefore not only could the number of potential tetraplexes at these loci be much larger, but the stability of these tetraplexes would be significantly higher as well.

A variety of other tandem repeats have been shown to form fold-back structures. These include the 5’-CAG-3’ repeat that is unstable in triplet expansion diseases such as Huntington’s disease and myotonic dystrophy (26, 28, 29, 31) and the centromeric satellite sequence (27). Other simple satellites such as the A + T-rich hypervariable sequence in the 3’ region of the human apolipoprotein B gene (66) also have the potential to form cruciforms and hairpins. Some G + C-rich repeats may also form other unusual DNA structures such as triplexes (33).

In the strand slippage models for the generation and evolution of tandem arrays, the nascent strand dissociates from the template, allowing the two strands to slip relative to one another. Successful priming from the slipped position results in a change in repeat number. Factors that favor strand dissociation over polymerization or that stabilize a slipped nascent strand-template complex would be expected to affect the frequency with which repeat units are added to or lost from the array. Blocks to DNA synthesis, such as those resulting from tetraplex formation, would be expected to increase the likelihood that strand slippage would occur. Since the strongest blocks to DNA synthesis are encountered at the 3’ end of such an array, these structures would account for the polarity observed for the gain and loss of repeat units from tandem arrays (12, 16, 67). In addition, since polymerase pause sites are known to be hotspots for nucleotide misinsertions (68), such blocks could also explain the clustering of point mutations at one end of the array (12, 16, 67).

One model that attempts to explain the large scale increase in repeat number seen in some tandem arrays invokes a long lived block to DNA synthesis that induces repeat strand slippage during replication (20). Tetraplexes make compelling candidates for this long lived block since they form strong, stable blocks to DNA synthesis under physiological conditions (24, 25, 34). We have shown that even very long hairpins are not effective barriers to DNA polymerase in our assay (see Ref. 47 and Woodford et al.), which suggests that sequences that are only able to form hairpins may not arrest DNA synthesis. This would be consistent with in vivo observations (69). However, both tetraplexes and hairpins may act to increase the frequency of successful strand slippage by stabilizing the strand slippage intermediate, thus increasing the likelihood that reinitiation of the polymerase would occur from the slipped position.

In addition, we would expect that the intramolecular tetraplex-forming tandem arrays are also likely to form intermolecular tetraplexes involving either one or three other DNA strands (70). Formation of such structures may facilitate synapsis of the DNA strands prior to crossing over during recombination. A combination of enhanced pausing at intrastrand tetraplexes, and enhanced synapsis between strands from different chromosomes or chromatids, may promote instability by facilitating strand switching.

It is possible that the formation of secondary structures in general may contribute to the generation and evolution of tandem arrays. In this regard, we would expect that the likelihood of structure formation would be affected by a variety of factors including the nature of the flanking sequences, the local chromatin structure, the transcriptional activity of a region, the rate of replication through the tandem array, the size of individual nucleotide pools, and whether or not the secondary structure-forming sequence is in the leading or lagging strand of DNA synthesis (71).

Acknowledgments—We thank Drs. Anthony Furman and Herbert Tabor for critical reading of this manuscript and for their advice and support.

REFERENCES

1. Willard, H. F. (1989) Genome 31, 737–744
2. Caskey, C. T., Pizzuti, A., Fu, Y. H., Fenwick, R. J., and Nelson, D. L. (1992) Science 256, 784–789
3. Nussbaum, R. L., and Ledbetter, D. H. (1986) Annu. Rev. Genet. 20, 109–145
4. Stephan, W., and Cho, S. (1994) Genetics 136, 333–341
5. Stephenson, W. (1989) Mol. Biol. Evol. 6, 188–212
6. Weber, J. L., and Wong, C. (1995) Hum. Mol. Genet. 2, 1123–1128
7. Willard, H. F. (1991)Curr. Opin. Genet. & Dev. 1, 509–514
8. Armour, J. A., Wong, Z., Wilson, V., Royle, N., and Jeffreys, A. J. (1989) Nucleic Acids Res. 17, 4925–4935
9. Armour, J. A., Harris, P. C., and Jeffreys, A. J. (1988) Hum. Mol. Genet. 2, 1137–1145
10. Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) Nature 314, 67–73
11. Jeffreys, A. J., Royle, N. J., Wilson, V., and Wong, Z. (1988) Nature 332, 278–281
12. Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., and Armour, J. A. (1994) Nat. Genet. 6, 136–145
13. Smith, G. P. (1974) Cold Spring Harbor Symp. Quant. Biol. 39, 507–513
14. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terasaki, E., and Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 77–84
15. Levinson, D., and Gutman, G. A. (1987) Mol. Biol. Evol. 4, 203–221
16. Kunst, C. B., and Warren, S. T. (1984) Cell 77, 853–861
17. Snow, K., Tester, D. J., Kruckeberg, K. E., Schaid, D. J., and Thibodeau, S. N. (1994) Hum. Mol. Genet. 3, 1543–1551
18. Wörheide, D., Hennig, I., Vogel, W., and Steinbach, P. (1993) Nat. Genet. 4, 140–142
19. Richards, R. I., and Sutherland, G. R. (1994) Nat. Genet. 6, 114–116
20. Wells, R. D., and Sinden, R. R. (1995) in Genomic Rearrangement and Stability (Davies, K. E., and Warren, S. T., eds) Vol. 7, 1st Ed., pp. 107–138, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. Jarman, A. P., and Wells, R. A. (1989) Trends Genet. 5, 367–371
22. Fu, Y., Kuhl, D. P., Pizzuti, A., Pietretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J., Holden, J. J., Fenwick, R. G., Warren, S. T., Ootstra, B. A., Nelson, D. L., and Caskey, C. T. (1993) Cell 77, 1047–1058
23. Verkerk, A. J., Pietretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., Reisner, O., Richards, S., Victoria, M. F., Zhang, F. P., Eussen, B. E., van Ommen, G. B., Blonden, L. A. J., Riggins, G. J., Chastain, J. L., Kunst, C. B., Galjaard, H., Caskey, C. T., Nelson, D. L., Oostra, B. A., and Warren, S. T. (1993) Cell 76, 905–914
24. Uus, K., and Woodford, K. J. (1995) Nucleic Acids Res. 23, 4202–4209
25. Weitzmann, M. N., Woodford, K. J., and Uus, K. (1996) J. Biol. Chem. 271, 29068–29074
26. Chen, X., Mairanpp, S. V., Castani, P., Ratliff, R., Mozay, R. K., Laazoun, Y., Smith, S. S., Bradbury, E. M., and Gupta, G. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 5199–5203
27. Ferrer, N., Azevin, F., Villasante, A., Gutierrez, C., and Abad, J. P. (1995) J. Mol. Biol. 245, 8–21
28. Cary, A. M., Goellner, G., Juranic, N., Macura, S., and McMurray, C. T. (1995) Cell 81, 533–540
29. Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., and Haworth, I. S. (1995) Nucleic Acids Res. 23, 1050–1059
30. Mitas, M., Yu, A., Dill, J., and Haworth, I. S. (1995) Biochemistry 34, 12603–12611
31. Mitchell, J. E., Newbury, S. F., and McCellan, J. A. (1995) Nucleic Acids Res. 23, 1876–1881
32. Yu, A., Dill, J., Wirth, S. S., Huang, G., Lee, V. H., Haworth, I. S., and Mitas, K. J. Woodford, M. N. Weitzmann, and K. Uusdin, unpublished observations.
33. Brereton, H. M., Fingaira, F. A., and Turner, D. R. (1993) *Nucleic Acids Res.* **21**, 2563–2569
34. Gelbert, M., Lipsett, M. N., and Davies, D. R. (1963) *Proc. Natl. Acad. Sci. U. S. A.* **48**, 2013–2018
35. Zimmerman, S. B., Cohen, G. H., and Davies, D. R. (1975) *J. Mol. Biol.* **92**, 181–192
36. Sen, D., and Gilbert, W. (1988) *Nature* **334**, 364–366
37. Fried, R., and Sinden, R. R. (1991) *Nature* **352**, 544–547
38. Arnesen, J. A., Crozier, M., Malcolm, S., Chan, J. C., and Jeffreys, A. J. (1995) *Proc. R. Soc. Lond. B Biol. Sci.* **261**, 345–349
39. Horn, G. T., McCleathay, A. I., Richards, B., MacDonald, M. E., and Gusella, J. F. (1991) *Nucleic Acids Res.* **19**, 4772
40. Baran, N., Lapidot, A., and Manor, H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 507–511
41. Hammond-Kosack, M. C., Debrinski, B., Lurz, R., Docherty, K., and Kilpatrick, M. W. (1992) *Nucleic Acids Res.* **20**, 231–236
42. Hammond-Kosack, M. C., Kilpatrick, M. W., and Docherty, K. (1993) *J. Mol. Endocrinol.* **10**, 121–126
43. Eichler, E. E., Hammond, H. A., Macpherson, J. N., Ward, P. A., and Nelson, D. L. (1995) *Hum. Mol. Genet.* **4**, 2199–2208
44. Trinh, T. Q., and Sinden, R. R. (1991) *Nature* **352**, 544–547
45. Kelly, R., Bulfield, M. R., Collick, A., Gibbs, M., and Jeffreys, A. J. (1989) *Genomics* **5**, 844–856
46. Segal, E. D., and Sinden, R. R. (1991) *Nature* **352**, 544–547
47. Witty, S., and Smith, E. (1989) *Nature* **344**, 410–414
48. Zhang, X., and Weintraub, H. (1991) *Nature* **352**, 544–547
49. Schierer, T., and Henderson, E. (1994) *Biochemistry* **33**, 2240–2246
50. Wang, L.-S., and Breslow, J. L. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7575–7579
51. Lapidot, A., and Manor, H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7575–7579
52. Schierer, T., and Henderson, E. (1994) *Biochemistry* **33**, 2240–2246
53. Hammond-Kosack, M. C., Debrinski, B., Lurz, R., Docherty, K., and Kilpatrick, M. W. (1992) *Cell Biol. Res.* **20**, 231–236
54. Hammond-Kosack, M. C., Kilpatrick, M. W., and Docherty, K. (1993) *J. Mol. Endocrinol.* **10**, 121–126
55. Ling, G. N. (ed) (1984) *In Search of the Physical Basis of Life*, pp. 227–269, Plenum Press, New York
56. Voloshin, O. N., Veselkov, A. G., Belotserkovskii, B. P., Danilevskaya, O. N., Pavlova, M. N., Dobrynin, V. N., and Frank-Kamenetskii, M. D. (1992) *J. Mol. Biol.* **9**, 645–652
DNA Secondary Structures and the Evolution of Hypervariable Tandem Arrays
Kerry J. Woodford, Karen Usdin and M. Neale Weitzmann

J. Biol. Chem. 1997, 272:9517-9523.
doi: 10.1074/jbc.272.14.9517

Access the most updated version of this article at http://www.jbc.org/content/272/14/9517

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 72 references, 21 of which can be accessed free at http://www.jbc.org/content/272/14/9517.full.html#ref-list-1