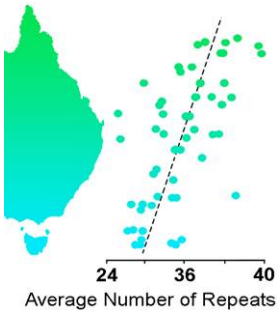
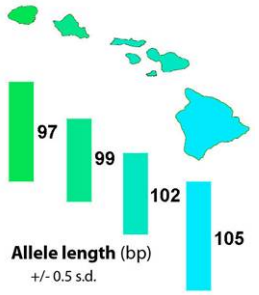
Facilitating Evolution:
Adaptation and Simple Sequence Repeat Polymorphism

David G. King[1] and Yechezkel Kashi[2]

[1]Dept. of Anatomy and Dept. of Zoology, Southern Illinois University Carbondale
[2]Dept. of Biotechnology and Food Engineering, The Technion - Israel Institute of Technology



In Australian *Drosophila*, average number of minisatellite repeats in a heat shock gene changes with latitude.
(after Collinge *et al.* 2008)



In Hawaiian mints, average length of a minisatellite in a gene for flowering time shifts with island age.
(after Lindqvist *et al.* 2007)

Background for **SSRs**
(see reference list on reverse side of this page)

SSRs (also called MICROSATELLITES and MINISATELLITES) are DNA tracts in which a relatively short base-pair sequence, or motif, is repeated over and over in tandem.

SSRs experience frequent, reversible, site-specific mutations which add or subtract motif units.

Many SSRs are located in functional domains, within exons, introns, and UTRs, as well as in upstream and downstream regulatory regions, where the number of motif repetitions can influence practically any aspect of gene function.

Background for **indirect selection**
(see reference list on reverse side of this page)

INDIRECT SELECTION of a trait which does not directly affect phenotype, such as site-specific mutability, occurs when the trait is closely and causally linked with beneficial mutations.

At SSR loci, mutability is encoded by motif length and repeat purity while phenotype is encoded by the number of repeats.

Repeat-number stability will be favored only when an ancestral number of repeats consistently yields higher fitness than do variants with altered repeat numbers.

Mutability at any given SSR locus will be selected (indirectly) when changing circumstances repeatedly favor mutant alleles.

Mutationally variable SSRs equip populations with advantageously adjustable genes, just as TUNING KNOBS confer beneficial adjustability on musical instruments.

Polymorphism in SIMPLE SEQUENCE REPEATS (SSRs) can influence the action of genes related to adaptation.

In *Drosophila*, repeat variation in the period gene and in the heat shock protein gene contributes to climate adaptation (Sawyer *et al.* 2006, Zamorzaeva *et al.* 2005; Collinge *et al.* 2008).

In *Saccharomyces*, variation in coding repeats in genes for cell-wall proteins creates variation in surface properties such as adhesion or biofilm formation (Verstrepen *et al.* 2005).

In pathogenic bacteria, SSR-based mutability in contingency genes confers adaptive advantage that can be experimentally demonstrated (Bayliss & Moxon 2006).

SSR mutability can thus enhance the responsiveness of populations to varying environmental conditions.

Although SSRs are often characterized as "junk", their exceptional mutability can contribute genetically meaningful variation.

When integrated into functional gene complexes, SSRs act as general-purpose **TUNING KNOBS** for adjusting gene function (Kashi *et al.* 1997, King *et al.* 1997, Kashi & King 2006a,b).

SSRs contribute to adaptation by providing an inexhaustible supply of quantitative genetic variation.

SSRs' site-specific, repeat-based mutability can facilitate efficient adaptive adjustment of quantitative traits.

Mutation rate and typical mutation size are based on heritable parameters of each SSR site, such as motif length and purity of motif repetition.

SSR mutability and location have plausibly been shaped by **INDIRECT SELECTION**, so that these sequences provide a reliable supply of low-cost genetic variation (King & Soller 1999, Kashi & King 2006b, King & Kashi 2007a,b).

Background for **SSR variation** *vs.* **SNPs**
(see reference list on reverse side of this page)

SSRs

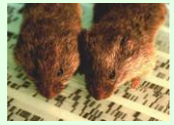
1. SSR repeat-number mutations occur at frequencies orders of magnitude greater than point mutations. Yet such mutation rates are not severely deleterious.
2. Variation in SSR mutation rate is governed by intrinsic attributes of each SSR site. A heritable, site-specific rate accompanies each SSR allele.
3. Phenotypic effects of SSR mutation are often quantitative and small, with fair likelihood of adaptive advantage. (Reversible switching also occurs.)
4. SSR mutations typically alter the number of repeating motifs. Such mutations can be readily reversed by subsequent mutation.

vs.
SNPs

1. Point mutations typically occur at extremely low rates, with natural selection apparently favoring the highest practical level of replication fidelity.
2. Variation in point mutation rate typically involves replication mechanisms whose effects are not site-specific but distributed throughout the genome.
3. Phenotypic effects of point mutations range from neutral to lethally disruptive, with beneficial effects occurring only rarely.
4. Multiple types of point mutation (substitutions, insertions, and deletions) make reversal of any particular mutation unlikely. Reversal of multiple mutations is highly improbable.

SSR variation affects morphology of dog breeds.
Image courtesy J. Fondon III

Some authors believe it to be as much the function of the reproductive system to produce individual differences, or very slight deviations of structure, as to make the child like its parents.
Charles Darwin, 1859



SSR variation affects vole social behavior.
Image courtesy L. Young

Facilitating Evolution: Adaptation and Simple Sequence Repeat Polymorphisms

*King, David G.; Kashi, Yechezkel

(DGK) Southern Illinois University Carbondale, Department of Anatomy and Department of Zoology, Carbondale, IL 62901; (YK) Israel Institute of Technology, Department of Biotechnology and Food Engineering, Haifa 32000, Israel. *e-mail* dgking@siu.edu

We highlight accumulating evidence that repeat-number polymorphism may play a causal role in ecological adaptation. Indeed, simple sequence repeats appear well-suited to provide an evolutionarily-significant source of adaptive genetic variation.

Simple sequence repeats (SSRs; microsatellites and minisatellites) are a rich source of genetic polymorphism. Although SSRs are commonly regarded as "junk", abundant evidence has demonstrated that repeat-number alleles are not all "neutral" but can exert a quantitative influence on practically any aspect of gene activity. A large proportion of all genes incorporate one or more SSRs in exons, in introns, in UTRs, and in upstream and downstream regulatory regions. The particular action of repeat-number variants at any such site is determined by the site's function and by the SSR's motif type. For example, the number of triplet repeats in a protein's coding region determines the length of a homopolymeric amino acid stretch, while variation in a non-triplet repeat within a coding sequence can effectively turn a gene on or off by shifting the reading frame. In UTRs, dinucleotide repeat variation can alter transcript shape and hence influence transcription activity. Repeat variation elsewhere can influence the binding of regulatory factors.

Innumerable studies have reported variation among natural populations in the frequency-distribution of SSR alleles. Several studies have also found a consistent correlation between an environmental parameter and the number of repeats within a gene whose function is related to that parameter. In at least a few such cases (e.g., the *period* gene in *Drosophila melanogaster*), experimental evidence indicates that the observed repeat-number variation does indeed influence gene function in a manner consistent with selection as a cause for the observed cline in allele frequencies. We argue that such effects are not just curious special cases but are examples of a principal of broad significance.

Indeed, indirect selection has plausibly shaped SSR mutability and location so that these sequences can provide a reliable supply of low-cost genetic variation. Indirect selection occurs when a heritable parameter such site-specific mutability, which does not directly affect phenotype, is closely and causally linked with beneficial mutant phenotypes. At SSR loci, mutability is encoded by motif length and repeat purity while phenotype is encoded by the number of repeats, so that the traits of mutation rate and typical mutation size for any SSR should both be amenable to indirect selection. Repeat-number stability will be favored only when an ancestral number of repeats consistently yields higher fitness than do variants with altered repeat numbers. Mutability will be selected (indirectly) when changing circumstances repeatedly favor mutant alleles.

Thus the site-specific, repeat-based mutability which characterizes SSRs may have been selected, indirectly, because it can facilitate efficient adaptive adjustment of quantitative traits.

Reviews

Kashi Y, King, DG and Soller M (1997) Simple sequence repeats as a source of quantitative genetic variation. *Trends in Genetics* 13: 74-78.

King DG, Soller M, and Kashi Y (1997) Evolutionary tuning knobs. *Endeavour* 21: 36-40.

Kashi Y and King DG (2006a) Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* 22: 253-259.

King DG, Trifonov EN and Kashi, Y (2006) Tuning knobs in the genome: Evolution of simple sequence repeats by indirect selection. In: LH Caporale, ed., *The Implicit Genome*, Oxford University Press.

Fondon III JW *et al*. (2008) Simple sequence repeats: Genetic modulators of brain function and behavior. *Trends in Neuroscience* 31: 328-334

Mutability and indirect selection

King, DG, and Soller, M (1999) Variation and fidelity: The evolution of simple sequence repeats as functional elements in adjustable genes. In: S.P. Wasser, ed., *Evolutionary Theory and Processes: Modern Perspectives*, Kluwer Academic Publ.

Kashi Y and King DG (2006b) Has simple sequence repeat mutability been selected to facilitate evolution? *Israel Journal of Ecology and Evolution* 52:331-342.

King DG, Trifonov EN and Kashi, Y (2006) Tuning knobs in the genome: Evolution of simple sequence repeats by indirect selection. In: LH Caporale, ed., *The Implicit Genome*, Oxford University Press.

King DG and Kashi Y (2007a) Mutability and Evolvability: Indirect selection for mutability. *Heredity* 99:123-124.

King DG and Kashi Y (2007b) Mutation rate variation in eukaryotes: evolutionary implications of site-specific mechanisms. *Nature Reviews Genetics* 8 (doi:10.1038/nrg2158-c1).

Population genetics and Drosophila temperature compensation

Zamorzaeva I *et al*. 2005 Sequence polymorphism of candidate behavioural genes in *Drosophila melanogaster* flies from 'Evolution Canyon'. *Molecular Ecology* **14**: 3235–3245.

Sawyer LA *et.al*. 2006 The *period* gene Thr-Gly polymorphism in Australian and African *Drosophila melanogaster* populations: Implications for selection. *Genetics* **174**: 465-480.

Collinge JE *et al*. 2008 Latitudinal and cold-tolerance variation associate with DNA repeat-number variation in the hsr-omega RNA gene of *Drosophila melanogaster*. *Heredity* 101:260–270

Adaptive radiation of Hawaiian endemic mints

Lindqvist C *et al*. 2007 Polyglutamine variation in a flowering time protein correlates with island age in a Hawaiian plant radiation. *BMC Evolutionary Biology* 2007, **7**:105 (doi:10.1186/1471-2148-7-105).

Bacterial pathogenicity

Bayliss CD and Moxon ER (2006) Repeats and variation in pathogen selection. In: LH Caporale, ed., *The Implicit Genome*, Oxford University Press.

Yeast ecology

Verstrepen KJ *et al*. (2005) Intragenic tandem repeats generate functional variability. *Nature Genetics* 37, 986–990

Canine and feline morphology

Fondon III JW and Garner, HR (2004) Molecular origins of rapid and continuous morphological evolution. *PNAS* 101(52): 18058-18063.

Sears KE *et al*. (2007) The correlated evolution of *Runx2* tandem repeats, transcriptional activity, and facial length in Carnivora. *Evolution & Development* 9:555-565.

Vole (and primate) **social behavior**

Hammock EAD and Young LJ (2005) Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* 308: 1630–1634.

Transcriptional regulation

Gerber H-P *et al*. (1994) Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science* 263: 808-811.