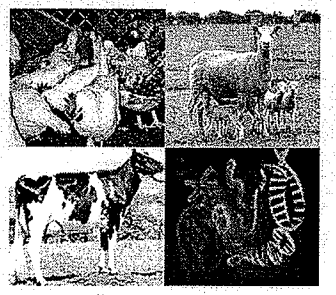
Molecular markers and maps



Karen Marshall

UNE
THE UNIVERSITY OF NEW ENGLAND

IAEA, Korea, April, 2006

Molecular markers

Are sites where differences in DNA sequences occur among members of the same species

'reveal polymorphisms at the DNA level'

can be in either coding or non-coding regions

Variations at the DNA level

Single nucleotide polymorphisms (SNPs)

Insertions or deletions (Indels)

Variable number of tandem repeats (VNTRs)

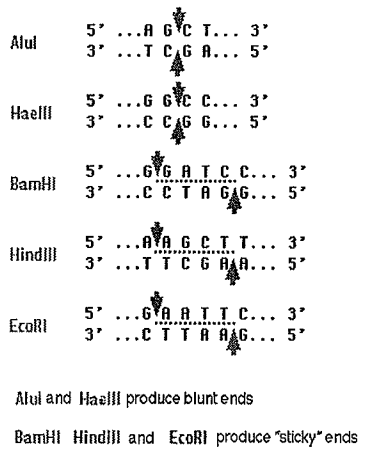
Markers detect one or more of these variations

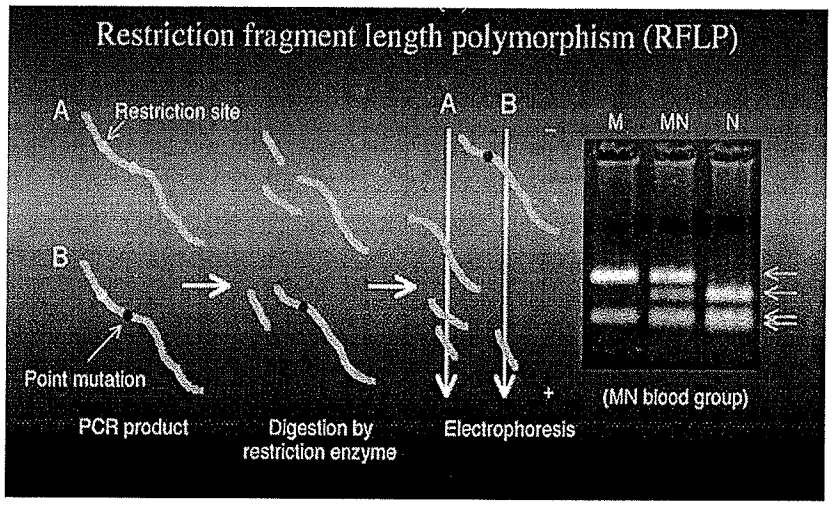
RFLP markers

Restriction fragment length polymorphisms

Restriction enzymes recognise and cut DNA at specific sites

Different sized fragments are produced depending on whether the restriction site exists or not





Markers and maps

Microsatellite markers

Type of VNTR, which are multiple copies of a sequence of base pairs arranged end to end

Length of repeating unit varies

- if <4 base pairs: microsatellite
- if >4 base pairs: minisatellite

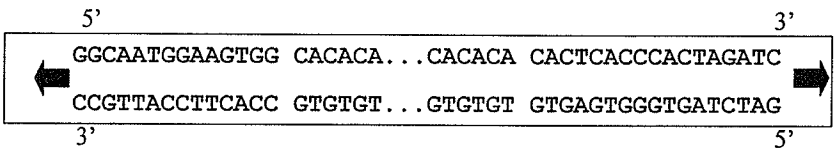
5′ CACACACACACA 3′
3′ GTGTGTGTGTGT 5′

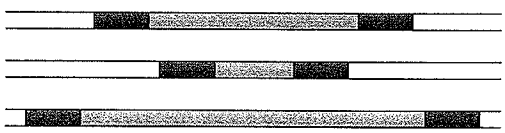
Also notated $(CA)_n$

Markers and maps

Microsatellite markers

BL25





Alleles differ in length

Markers and maps

Typing microsatellites

Most commonly use PCR based methods

Steps are

- amplify region by PCR
  - *primers labelled via radioactivity or fluorescence*
- separate PCR products according to size
  - *polyacrylamide gel, capillary based systems*
- determine size of amplified product
  - *autoradiography, fluorescent traces*
- score alleles

Markers and maps

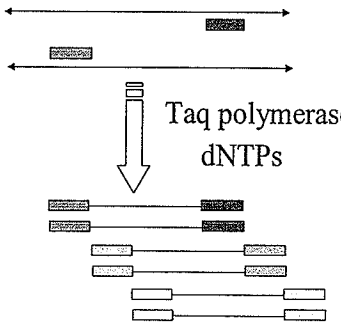
PCR

Denaturation

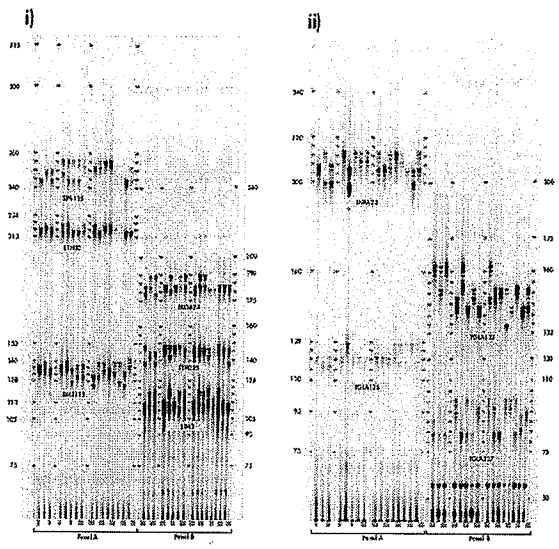
Annealing

Extension

~25 cycles



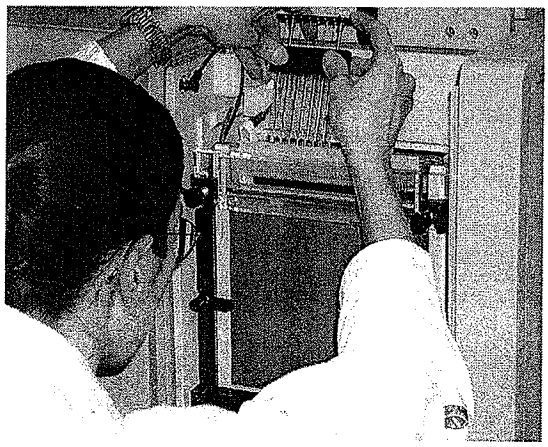
Markers and maps



Bovine Microsatellite Multiplexing for Herd Evaluation and Parentage

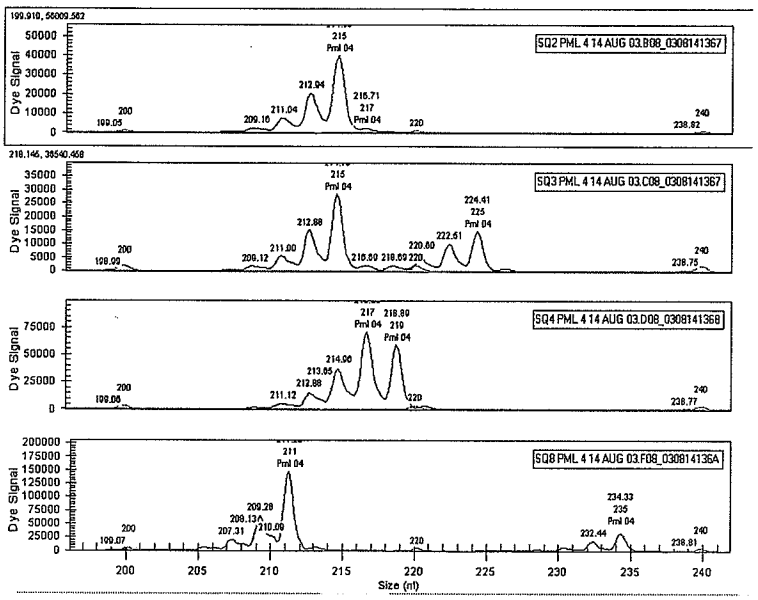
J. Kovar, J. Walker, D. Steffens, J. Harford, and J. Qiu
LI-COR Inc., 4308 Progressive Avenue, Lincoln, NE 68504, USA

Markers and maps



ABI 377

Markers and maps



Markers and maps

SNP markers

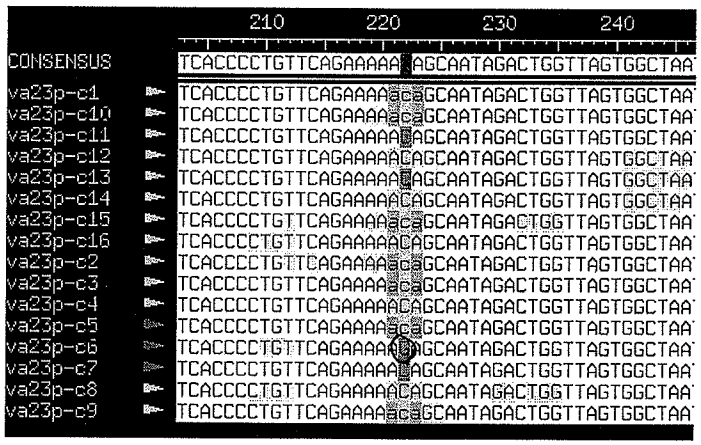
Single base change in DNA sequence

Usually two alternative nucleotides at a single position

Least frequent allele present at 1% or greater

Why not 4 alternative nucleotides?

- low prob. of 2 independent base change occuring at any single position
  - (1-5 x 10^{-9} / nucleotide / generation at neutral position)
- Base for transitional mutations (A ⇔ G, C⇔ T) over transversions

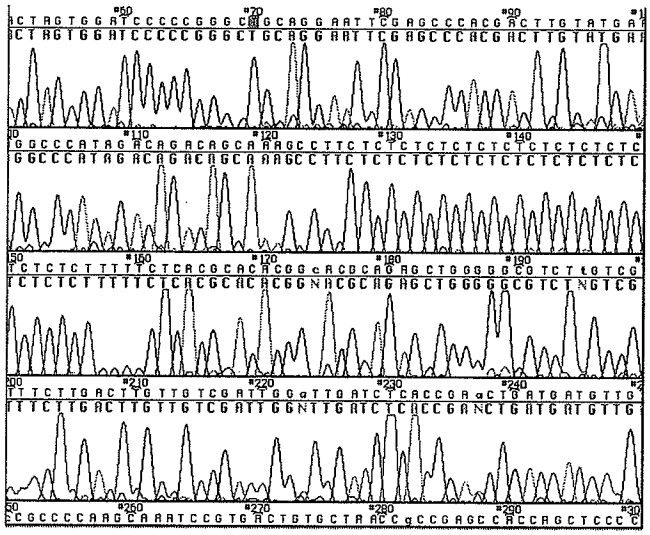


Typing SNPs

Numerous methods

- Direct sequencing
- DNA chips (potential for very high throughput)
- Other e.g. SSCP (single stranded confirmation polymorphism), primer extension, pyrosequencing etc. (see Vigal et al. GSE 2002)

Markers and maps



Markers and maps

Variations detected by markers

Marker	Variation type		
	SNP	Indel	VNTR
RFLP	+	(+)	(+)
Microsatellite	-	(+)	+
SNP	+	(+)	-
RAPD (random amplification of polymorphic DNA)	+	(+)	(+)
AFLP (amplified fragment length polymorphism)	+	(+)	(+)
SSCP (single stranded confirmation polymorphism)	+	(+)	(+)

From: Vignal et al. GSE 2002.

Markers and maps

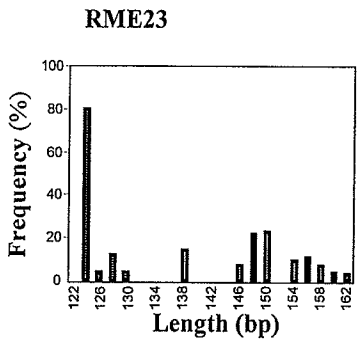
Properties of markers: statistical considerations

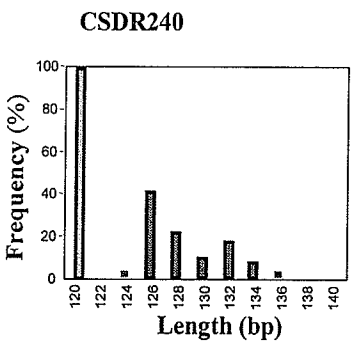
Heterozygosity

- SNPs: two co-dominant alleles
- microsatellites: numerous co-dominant alleles
- thus, lower heterozygosity of single locus SNPS compared to microsatellites
- note, however, that marker heterozygosity is always population dependent

Markers and maps

Microsatellite allele frequency





Merino

Bos taurus x Bos indicus, N'Dama, Boran, Brangus

Brahman, Hereford, Africander

Merino, Suffolk, Border Liester, Romney, Poll Dorset & Others

Properties of markers: statistical considerations

Density
- SNPs (~1 every 1000 bp)>> microsatellites

Neutrality
- imp. assumption of pop'n genetics
- microsatellites usually in con-coding regions, whereas neutrality of SNPs is case dependent

Mutation rate
- Microsatellites ($1x10^{-5}$) > SNPs ($1x10^{-9}$)

Rate and type of genotyping errors

To trace livestock domestication

Markers may be located on

- Mitochondrial DNA
  - maternal inheritance
  - choice for domestication studies
- Y chromosome
  - paternal inheritance
  - for many livestock species few polymorphic Y markers exist (may indicate small # of male lineages)
- Autosomes
  - bi-parental inheritance

Markers and maps

Maps

Genetic map

- Order and location of markers assigned to chromosome on the basis of linkage analysis
- Distance measured in Morgans (M)

Physical map

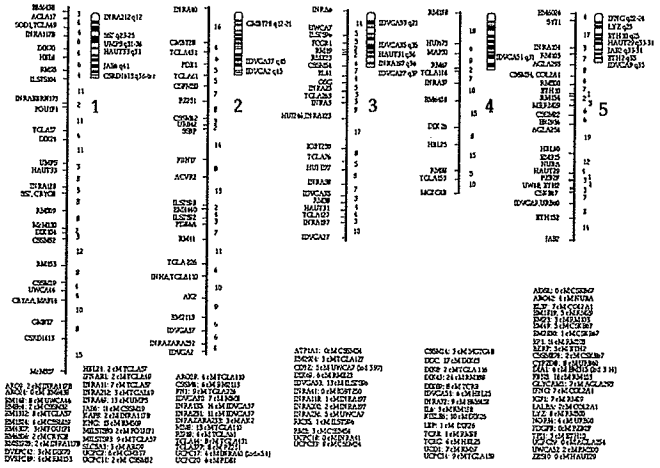
- Actual structure of genetic material
- At highest level DNA sequence
- Distance measured in 10^6bp (Mbp)

Genetic and physical maps are usually 'linked' together

Markers and maps

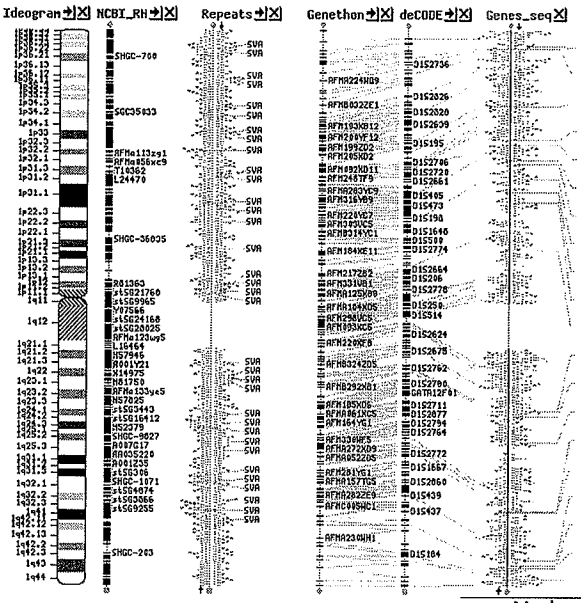
Example of a map

Cattle chromosomes 1-5, from Cattle Genome Database hosted at the Queensland Biosciences Precinct: http://www.cgd.csiro.au



Markers and maps

Further example - Human Chr 1



Markers and maps

Genetic maps

Map distance is determined from the number of observed recombination events

1cM = 100 recombination event per 100 meiosis

Only odd number of cross-over events are observed

- r = probability of an odd-number of cross-over events
- 1-r = probability of an even number of cross-over events, including zero

Markers and maps

Mapping functions

Mapping functions predict number of cross-over events from observed recombination events

Haldane (1919)

- assumes no-interferance (cross-overs occur randomly and independantly over the entire chromosome)
- M = -(ln(1-2r))/2

Kosambi (1944)

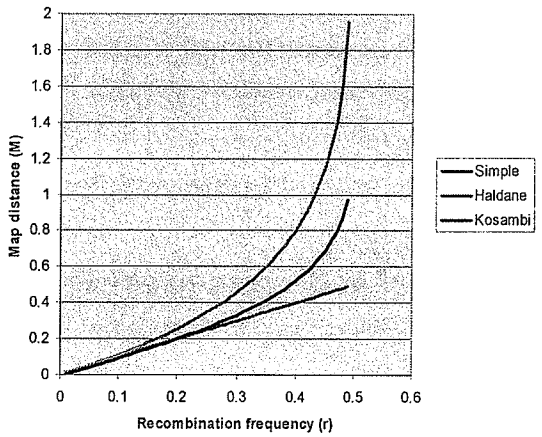
- assumes moderate interferance (i.e. some cross-over interferance at adjacent sites)
- M = ¼ ln (1+2r/1-2r)

Simple

- assumes complete interferances
- r = M

Markers and maps

Comparison of mapping functions



Markers and maps

Constructing linkage maps

Identification of recombinant gametes is easier if

- linkage phase of parents is known
  - sire AaBb x Dam AABB → 9 AABB, 1 AaBB, 1 AABb, 9 AaBb
  - sire thus gave gametes in frequency 0.45 AB, 0.05 aB, 0.05 Ab, 0.45ab: most likely phase is AB ab
- haplotype of gamettes transmitted from parents to offspring is known
  - AaBb x AABB → AaBb, sire gave ab dam gave AB
  - AaBb x AaBb → AaBb, cannot determine transmitted haplotypes

ML methods used to determine most likely phase

Markers and maps

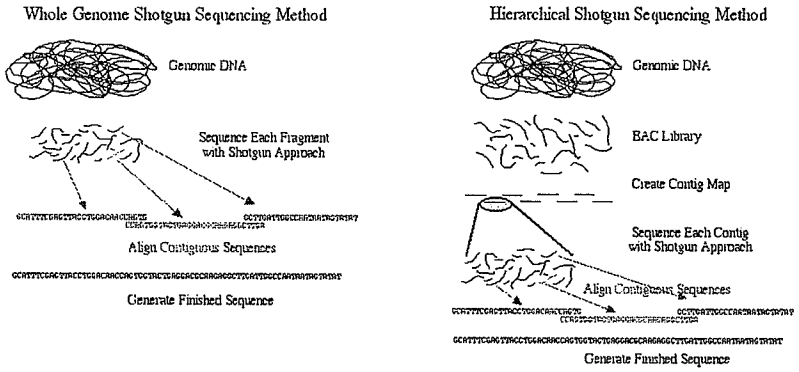
Physical maps

Various types

- Cytogenetic maps
  - banding pattern observed under light microscopy of stained chromosomes
  - low resolution (only estimates of the number of bp)
- Radiation hybrid
  - Use breaks induced by radiation to determine the distance between two markers
- Sequence tag sites (STS)
  - STS are short (100-500bp), unique DNA sequences with known location, can be derived from ESTs
- Sequence maps
  - 'the ultimate'

Cytogenic map - mouse

Generation of sequence maps



Relationship between genetic and physical distance

No universal relationship

- comparison of human genetic and sequence based physical maps, Yu et al. "Recombination rates varied greatly along each chromosome, from 0 to at least 9 centiMorgans per megabase"

Various depending on

- species
- chromosomal region: cross-overs often suppressed at centromeres, telomeres
- sex: female mammals usually have greater map distances than males, no crossing over in male Drosophila

Maps for livestock species

NCBI http://www.ncbi.nlm.nih.gov/mapview/

Cattle

In June 2005, the Human Genome Sequencing Center at Baylor College of Medicine released a 6.2X WGS assembly of the bovine genome (Btau_2.0). The source of the DNA was a female of the Hereford breed. The NCBI *Bos taurus* build 2.1 includes the 6.2X WGS Btau_2.0 assembly and a complete mitochondrial genome derived from a Korean native cow. The bovine genome is organized in 29 pairs of autosomes and a pair of sex chromosomes, X and Y.

Chicken

In March 2004 the Genome Sequencing Center at the Washington University School of Medicine in St. Louis released the assembled chicken genome. The chicken genome, the first avian genome to be sequenced, has a haploid genome size of 1200 Mb. The chicken genome, similar to other avian genomes, is composed of chromosomes of vastly different sizes identified as either macro- or microchromosomes. The *Gallus gallus* genome has 38 pairs of autosomes and a pair of sex chromosomes referred to as Z and W to distinguish them from mammalian sex chromosomes. In birds, the male is homozygous (ZZ) while the female is heterozygous (ZW).

Markers and maps

Maps for livestock species

NCBI http://www.ncbi.nlm.nih.gov/mapview

Sheep

The NCBI Map Viewer presents two genetic maps, (SM4.2) and (CAB), for *Ovis aries*.

The SM4.2 (SheepMap4.2) comprehensive linkage map has been provided by Dr. Jill Maddox (Centre for Animal Biotechnology, University of Melbourne, Australia). The SM4.2 map was produced on 11th June 2003 and represents an expansion of the SM3 map described in (Maddox et al., 2001).
SM4.2 comprises 1,232 loci and spans ~3,630 cM. This corresponds to almost complete coverage of the sheep genome. Each chromosome is represented by a single linkage group, with the largest gap between adjacent loci being 19.8 cM. This map was developed by genotyping the International Mapping Flock (IMF). The IMF was produced by AgResearch (NewZealand) in (Crawford et al., 1995).

The Meat Animal Research Center (MARC) map (CAB), is a genetic map kindly provided by Dr. John Keele. The CAB map was initially described in (de Gortari et al., 1998). The CAB map comprises over 500 markers and spans ~3063 cM.

Pigs

The NCBI Map Viewer presents a graphical view of the MARC linkage map for pig.

Markers and maps

Perspective
Advances in livestock genomics: Opening the barn door
James E. Womack

Department of Veterinary Pathobiology, Center for Animal Biotechnology and Genomics, Texas A&M University, College Station, Texas 77843-4467, USA

Genome research in animals used in agriculture has progressed rapidly in recent years, moving from rudimentary genome maps to trait maps to gene discovery. These advances are the result of animal genome projects following closely in the footsteps of the Human Genome Project, which has opened the door to genome research in farm animals. In return, genome research in livestock species is contributing to our understanding of chromosome evolution and to informing the human genome. Enhancement of these contributions plus the much anticipated application of DNA-based tools to animal health and production can be expected as livestock genomics enters its sequencing era.

Genome Research 15:1699-1705, 2005

Markers and maps