Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 951

Evolutionary Studies of the Mammalian Y Chromosome

BY

LINDA HELLBORG

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2004

Dissertation presented at Uppsala University to be publicly examined in Zootissalen, EBC, Uppsala, Friday, April 16, 2004 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Hellborg, L. 2004. Evolutionary Studies of the Mammalian Y Chromosome. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 951. 55 pp. Uppsala. ISBN 91-554-5904-8

Sex chromosomes are useful in elucidating the evolutionary factors affecting diversity and divergence. In particular, Y chromosome analyses may complement studies using mitochondrial DNA for inferring sex-specific population genetic processes.

Y chromosome studies have been scarce due to limited access to genetic markers and the dynamic evolution of Y. Conserved Y-specific primers that could amplify a diverse set of mammalian species were developed from comparison of gametologous X and Y sequences. Y-specific sequence, generally more than one kb, was amplified for all 20 species examined.

Intraspecific diversity on mammalian Y was found to be reduced even when male-biased mutation rate and effective population size were corrected for. A number of factors can cause this low variation on Y of which selection on a haploid chromosome seems most important.

The field vole (*Microtus agrestis*), a common and well-studied small mammal in Eurasia, was examined for X and Y variability. Earlier studies on mtDNA had shown that the field vole is separated in two distinct lineages in Europe. The X and Y chromosome sequences confirmed the deep split and suggested that the two lineages of field vole should be reclassified as two separate species.

Two distinct Y chromosome haplogroups were found in modern European cattle, distributed among breeds according to a north-south gradient. Ancient DNA analysis of European aurochsen showed the northern haplogroup to be the most common, possibly indicating local hybridization between domestic cows and wild aurochs bulls in Europe.

Keywords: Y chromosome, field vole, cattle, selection, primer design

Linda Hellborg, Department of Evolutionary Biology, Norbyv. 18 D, Uppsala University, SE-75236 Uppsala, Sweden

ISSN 1104-232X
ISBN 91-554-5904-8
urn:nbn:se:uu:diva-4126 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4126)

For Mattias & Eleonor

List of Papers

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals. Papers I and II are reprinted with permission from the publishers.

I Hellborg, L. and Ellegren, H. (2003)
Y Chromosome Conserved Anchored Tagged Sequences (YCATS) for the Analysis of Mammalian Male-specific DNA. *Mol Ecol* 12(1):283-291

II Hellborg, L. and Ellegren, H. (2004)
Low Levels of Nucleotide Diversity in Mammalian Y Chromosomes. *Mol Biol Evol* 21(1):158-163

III Hellborg, L., Gündüz, I. and Jaarola, M.
Sex-linked Markers Propose a New Mammalian Species in Europe. (Submitted)

IV Götherström, A., Hellborg, L., Anderung, C., Elburg, R., Smith, C. and Ellegren, H.
Early Cattle Husbandry: Y Chromosome Haplogroups Indicate Aurochs Introgression in Europe. (Manuscript)

Additional papers not included in the thesis:

Spong, G., Hellborg, L and Creel, S. (2000)
Sex Ratio of Leopards Taken in Trophy Hunting: Genetic Data from Tanzania. *Cons Gen* 1: 169-171

Hellborg, L., Walker, C.W., Rueness E.K., Stacy, J.E., Kojola, I., Valdmann, H., Vilá, C. Zimmermann, B., Jakobsen, K.S. and Ellegren, H. (2002)
Differentiation and Levels of Genetic Variation in Northern European Lynx (*Lynx lynx*) Populations revealed by microsatellites and mitochondrial DNA analysis. *Cons Gen* 3: 97-111

Spong, G. and Hellborg, L. (2002)
A Near-Extinction Event in Lynx: Do Microsatellite Data Tell the Tale? *Cons Ecol* 6: 15-19

Rueness, E.K., Jorde, P.E., Hellborg, L., Stenseth, N.C., Ellegren, H. and Jakobsen, K.S. (2003)
Cryptic Population Structure in a Large, Mobile Mammalian Predator: the Scandinavian Lynx. *Mol Ecol* 12: 2623-2633

Arrendal, J., Walker, C.W., Sundqvist, A-K., Hellborg, L. and Vilà, C. (In Press)
Genetic evaluation of an otter translocation program (*Conservation Genetics*)

Lindgren, G., Backström, N., Swinburne, J., Hellborg, L., Einarsson, A., Sandberg, K., Cothran, G., Vilà, C., Binns, M. & Ellegren H. (In Press)
Limited number of patrilines in horse domestication. (*Nature Genetics*)

Contents

Abbreviations

AFLP	Amplified fragment length polymorphism
AZF	Azoospermia factor
DNA	Deoxyribonucleic acid
MSY	Male specific Y
mtDNA	Mitochondrial DNA
Mya	Million years ago
PAR	Pseudoautosomal region
RAPD	Random amplified polymorphic DNA
RRS	Reduced representation shotgun
SNP	Single nucleotide polymorphism
SRY	Sex determining region of the Y
STR	Short tandem repeat
XAR	X added region
XCR	X conserved region
YCATS	Y chromosome conserved anchored tagged sequence

Why Y?

The Y chromosome is a gene poor, male specific chromosome in species with male heterogamety like humans. It often determines sex in a dominant fashion and is inherited clonally from father to son, so it is never present in females. In many higher organisms (with an X-Y sex determining system), it is the only chromosome with truly haploid characteristics where no material is exchanged with a homologue through recombination making all sites linked to each other. This non-recombining region is called the male-specific region of the Y chromosome, the MSY, and comprises 95% of the length of the chromosome in humans. The 5% that is genetically similar to X make up the pseudoautosomal region, PAR, in the telomere ends of the Y chromosome and recombine with the X chromosome during meioses. These unique features make Y very interesting and useful in a wide range of areas, including evolutionary and population genetic studies and forensic analyses. In this thesis, development of Y chromosome markers and their application will be presented and discussed. First, some background to the development of sex and the Y chromosome.

Sex determination

Sex determination mechanisms have evolved many times among eukaryotes and include a variety of environmental and genetic systems. The evolution of sex itself represents one of the major questions in evolutionary biology. Advantages include, for example, an increased opportunity to adapt to environmental changes due to favorable recombinant types. Recombination can also bunch together several deleterious alleles and eliminate them simultaneously (Crow 1994; Maynard-Smith 1978). Sexual populations can have a more rapid rate of evolution than an equivalent group of asexual

organisms (Barton and Charlesworth 1998; Crow and Kimura 1965). Asexual reproduction on the other hand produces twice the number of offspring compared to a sexually reproducing individual. Hermaphrodites (being both male and female at the same time) combine the two systems of asexual and sexual reproduction.

Organisms with male heterogamety, like mammals, are characterized by XY sex chromosomes in males and XX sex chromosomes in females. Sex chromosome aneuploidi, in humans and many other mammals, show that the Y chromosome determines sex since XXY karyotypes provide male phenotype and X0 female phenotype (Jacobs and Strongs 1959). The gene that determines sex in most mammals is *SRY* (sex determining region of the Y).

In mammals, the X chromosome is large and gene rich, while the Y chromosome differs in size, from almost the same size as the X in monotremes, to completely gone in two species of mole vole (*Ellobius*) (Graves 1995; Just et al. 1995)

In snakes, birds and some amphibians, females have a ZW karyotype while males have a ZZ. Different lineages in birds show genetically almost identical Z chromosomes containing about 6-11% of the genome (Ohno 1967b; Shetty et al. 1999). The size of the avian W chromosome varies among orders, being almost as large as the Z chromosome in ratites to a very tiny chromosome in, for example, chicken (Marshall Graves and Shetty 2001; Ohno 1967b).

Alligators and turtles represent groups of organisms that have temperature dependent sex determination, where the incubation temperature around the eggs has a strong influence on offspring sex (Kalthoff 1996).

Sex chromosome evolution

The evolution of sex chromosomes from a pair of autosomes involves a series of steps (Figure 1). Consider a sex-determining locus being located on a proto-sex chromosome. Genes linked to the sex-determining locus may acquire sexually antagonistic functions. Selection will then favor a reduced recombination rate between the sex-determining locus and sexually

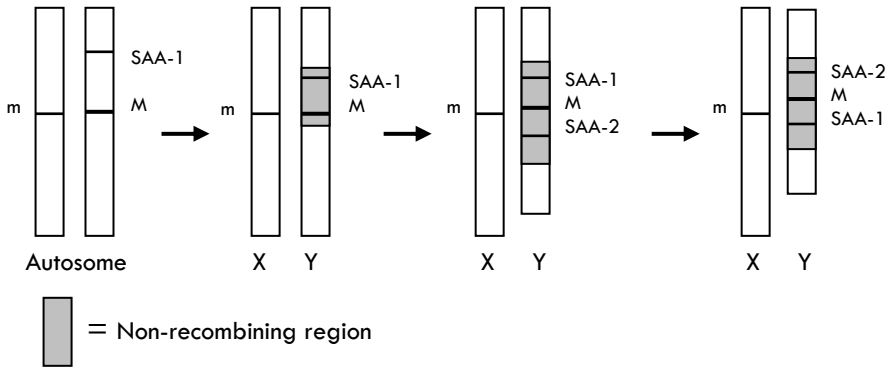

Figure 1. *Initiation of sex chromosome evolution from a pair of autosomes. A sex determining locus includes an allele leading to male development, M. At a linked locus a sexually antagonistic allele (SAA) favourable to males but not to females may arise. Selection will favour inhibition of recombination to make "male" genes be inherited together.*

antagonistic genes, as this will make sex-specific alleles be inherited together (Lahn and Page 1999a). In the absence of recombination, mutations will accumulate rapidly and degeneration of the non recombining region will occur fast due to mechanisms like "Muller's ratchet" (Figure 2) (Charlesworth and Charlesworth 2000). Muller's ratchet is the stochastic elimination of the chromosome class with the fewest deleterious mutations in the population. In the absence of recombination and back mutation this "least loaded" class can never be restored when lost (Felsenstein 1974; Muller 1964).

Degradation of the Y chromosome requires dosage compensation between the sexes, because dosage differences might be deleterious for many genes. Marsupials and most placental mammals inactivate the X chromosome genes with no Y-linked gametologue (sex chromosome genes with a common origin that have ceased to recombine) in the homogametic sex. This phenomenon is evident from a "barr body" seen in cytogenetic preparations in females, featuring the inactivated X chromosome. The particular X chromosome being inactivated varies between cells and all their mitotic descendants will have the same X inactivated. A female with a heterozygous

sex linked trait will express the two different alleles in patches throughout her body, which is evident in calico cats (Fagan 1979) (Figure 3).

In the eutherian lineage a translocation from an autosome to the undifferentiated region of X and Y occurred after the split from marsupials and before the main radiation of mammalian orders and is referred to as the X added region (XAR) (Figure 4). Many of the genes in this region do not undergo X inactivation. The original genes present on the proto sex chromosomes in all mammals, including monotremes and marsupials, are called the X conserved region (XCR) (Graves 1995).

Directed genetic exchange (ectopic gene conversion) has been detected between the sex chromosomes. A piece of Y chromosome sequence is replaced by a homologous sequence from X, which has been shown between the *Zfx/Zfy* gene pair (Hayashida et al. 1992; Pecon Slattery et al. 2000). Ectopic gene conversion may be a mechanism for adaptive correction of Y genes whose X homologues do not undergo X inactivation in females.

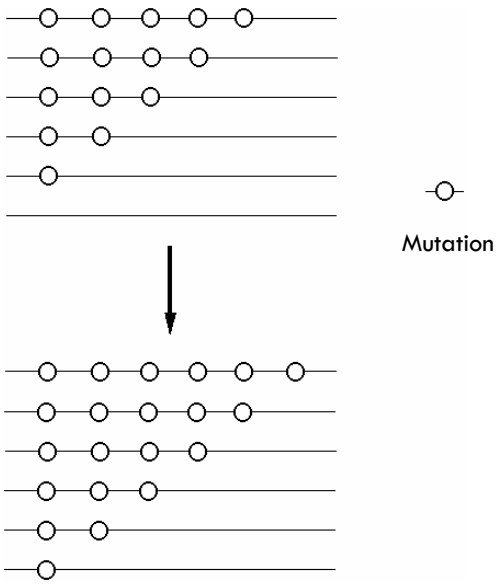

Figure 2. *Muller's ratchet. For each "turn" indicated by the arrow, the least loaded class disappears and can then not be restored.*

Figure 3.
A calico cat

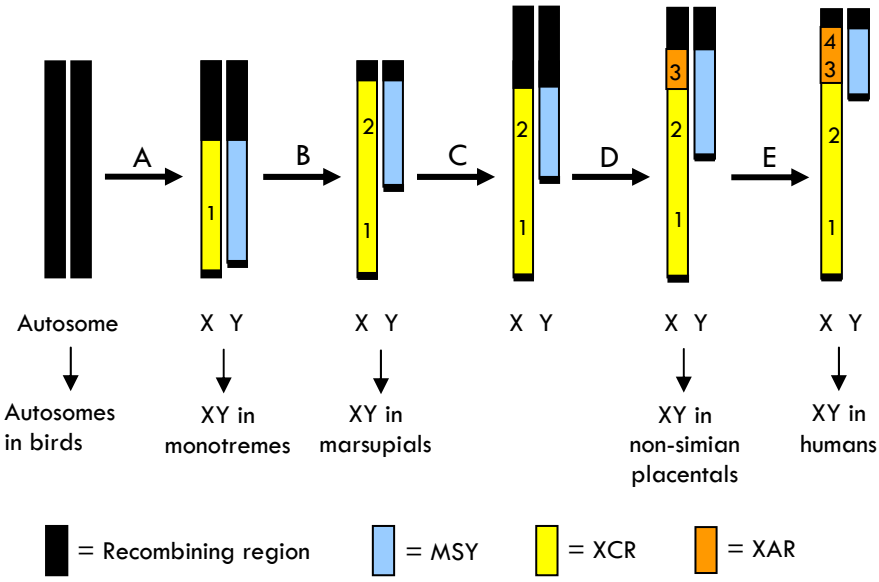

Figure 4. *Sex chromosome evolution in mammals: **A**) Male specific Y region (MSY) emerges by cessation of recombination with the X chromosome in a region including the SRY ≈230-350 Myr ago, **B**) Expansion of the MSY ≈130-170 Myr ago, **C**) Translocation expands the PAR ≈80-130 Myr ago, **D**) Further expansion of the MSY ≈80-130 Myr ago, **E**) Yet another expansion of the MSY ≈30-50 Myr ago. The expansion of MSY was probably accompanied by inversions preventing recombination. Numbers on the X chromosome refers to evolutionary strata .Modified from Lahn and Page (1999a)*

Evolutionary strata on the X chromosome

Cessation of recombination between the proto-X and Y chromosome was a necessary step towards their differentiation (Lahn and Page 1999a; Sandstedt

and Tucker 2004). In the human lineage, this occurred in at least four steps. A neutral estimate of divergence between X and Y gametologous can be obtained by calculating the synonymous substitution rate, Ks, between them. Lahn and Page used such Ks estimates to provide a measure of evolutionary time that has passed since recombination ceased in particular regions ("strata"); the higher Ks (the more divergent X and Y) the older the stratum. Four distinct groups were found with respect to Ks values and these were arranged in an orderly sequence on the X chromosome, from Xq to Xp. Estimated divergence times are for stratum 1 (240-320 Mya) followed by stratum 2 (130-170 Mya), stratum 3 (80-130 Mya) and stratum 4 where differentiation began when the New and Old world monkey lineages diverged (at least 30 Mya) (Figure 4). Strata 1 and 2 correspond to the original proto-X chromosome present in all mammals (called XCR by Graves), accordingly strata 3 and 4 correspond to the XAR only present in placental mammals (Graves 1995). A corresponding relationship was not found on Y most likely due to chromosomal inversions on the Y chromosome.

Evolution of the Y chromosome

Most of the genes on the Y chromosome have homologs on the X (Table 1 and Figure 5). However, there are some multicopied, testis-specific genes that appear to have been acquired by transposition or translocation from other parts in the genome (Skaletsky et al. 2003). For instance, *DAZ* was transposed from *DAZL* on chromosome 3 (Kuroda-Kawaguchi et al. 2001) and *CDY* was retrotransposed (by a processed messenger RNA) from a gene on chromosome 13 (Lahn and Page 1999b).

A comparison of orthologous Y chromosome genes in different species reveals a spectrum of functions and various stages of degeneration. One example is *ZFY* that in humans is a single copy gene with an X linked homologue where both the X and Y copy are expressed ubiquitously in the body. However, in rodents *ZFY* is a multicopy gene with testis-specific function (Koopman et al. 1991). Another example is *UBE1Y*, which is a putative spermatogenesis gene in marsupials and mouse but is lacking in primates. Perhaps it was lost in the primate lineage after its function in

spermatogenesis was taken over by an autosomal ubiquitin activating enzyme (Mitchell et al. 1998).

The fate of Y chromosome genes seems to follow a predetermined path when recombination ceases, from active or partially active to degraded and inactive pseudogene and, finally, to complete loss. This scenario is likely to have been followed by >1400 X-specific genes whose homologues have disappeared from the Y (Marshall Graves 2002a) (Figure 6**)**.

The human Y chromosome has only four genes left from the original 1000 genes on the X conserved region and about 20 genes from the original 500 genes on the X added region (Marshall Graves 2002a; Marshall Graves 2002b).

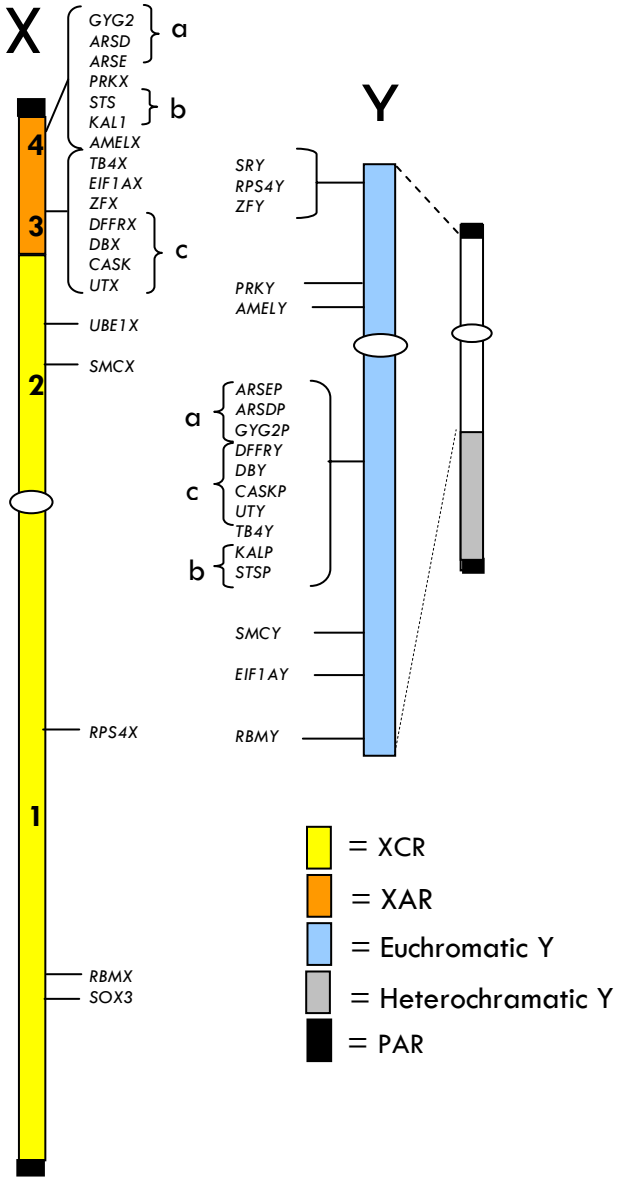

Figure 5. *Gametologous genes on the X and Y chromosome and their relationship to evolutionary strata (1-4 on the X chromosome), the X conserved region and X added region. Evidence of inversions on the Y chromosome can be seen by comparing the gene order on X and Y, especially the three small gene clusters a,b and c. Modified from Lahn and Page (1999a)*

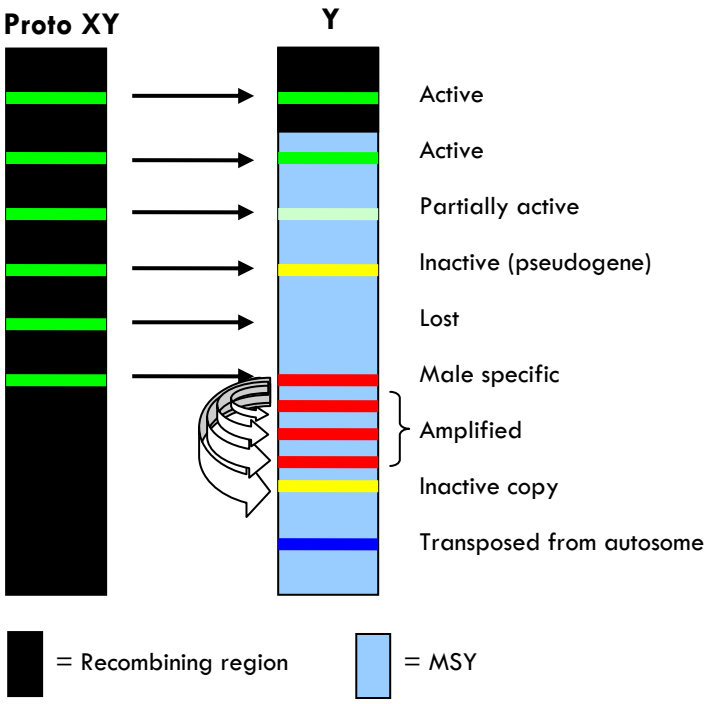

***Figure 6**. Evolution of genes on the Y chromosome from an active copy on the proto-XY chromosome to unchanged in the PAR of the Y. As Y degrades the gene can remain active (green), become partially active (pale green), become an inactive pseudogene or be completely lost. Some genes acquire male-specific functions (red), these are frequently amplified, and many copies are inactivated. A few genes are transposed from other parts of the genome. Modified from Marshall Graves (2002a).*

Structure of the human Y chromosome

The MSY is about 60 Mb in humans and consists of heterochromatic sequences (densely packed and never transcribed DNA) and three classes of euchromatic sequences: X-transposed, X-degenerate and ampliconic (palindrome regions). Over 150 transcription units have been found but only 78 seem to encode proteins (Skaletsky et al. 2003).

The X-transposed sequence is a 3.4 Mb transposed region from the long arm of the X chromosome that exhibits 99% identity to its origin. The transposition took place 3-4 million years ago in the human lineage and distinguishes humans from chimpanzees. The region encompasses only two genes, one testis specific (*TGIF2LY*) and one expressed in the brain (*PCDH11Y*).

The X-degenerate sequences are remnants of the ancestral pair of autosomes from which the X and Y chromosomes evolved. Together these sequences are 8.6 Mb in length and encompass 27 different single copy genes that are divided in 13 pseudogenes and 14 transcribed functional genes which have between 60% and 96% nucleotide sequence identity to their homolog on the X chromosome. The X-degenerate sequences encode 16 of the MSYs 27 distinct proteins or protein families. Most of the X-degenerate genes are expressed widely in the body except for *SRY*, *AMELY*, *TBL1Y* and *NLGN4Y* (Table 1).

The ampliconic class includes large repeated units, where sequences show greater than 99.9% identity to each other. The homology is thought to be maintained by frequent gene conversion (non-reciprocal transfer). The most prominent features here are eight massive palindromes, at least six of which contain testis genes. Nine different protein-coding gene families are found in this class with copy numbers from two to approximately 35 (Table 1). These genes are expressed predominantly or exclusively in the testis, probably with a function in spermatogenesis (Rozen et al. 2003).

The X-degenerate and the ampliconic classes dominate the MSY with 38% and 45% of the euchromatic sequences respectively. They are physically intermingled and neighboring genes display comparable diversities of male-specific ages. This implies a parallel evolution of the two classes as parts of a single molecule since the divergence from reptiles, 300 million years ago.

Table 1 MSY genes their copy numbers, expression and homologues (from Skaletsky et al. (2003)).

MSY sequence class	Gene symbol	Gene name	Number of copies	Tissue expression	X-linked homologue	Autosomal homologue
X-transposed	*TGIF2LY*	TGF (beta)-induced transcription factor	1	Testis	*TGIF2LX*	
	PCDH11Y	Protocadherin 11 Y	1	Fetal brain, brain	*PCDH11X*	
X-degenerate	*SRY*	Sex detrmining regionY	1	Predominantly testis	*SOX*	
	RPS4Y1	Ribosomal protein S4Y	1	Ubiquitous	*RPS4X*	
	ZFY	Zink finger Y	1	Ubiquitous	*ZFX*	
	AMELY	Amelogenin Y	1	Teeth	*AMELX*	
	TBL1Y	Transducin (beta)-like 1 protein Y	1	Fetal brain, prostate	*TBL1X*	
	PRKY	Protein kinase Y	1	Ubiquitous	*PRKX*	
	USP9Y	Ubiquitin-specific protease 9 Y	1	Ubiquitous	*USP9X*	
	DBY	Dead box Y	1	Ubiquitous	*DBX*	
	UTY	Ubiquitous TPR motif Y	1	Ubiquitous	*UTX*	
	TMSB4Y	Thymosin (beta)-4 Y	1	Ubiquitous	*TMSB4X*	
	NLGN4Y	Neuroligin 4 isoform Y	1	Fetal brain, brain, prostate, testis	*NLGN4X*	
	CYorf15A	Chromosome Y open reading frame 15A	1	Ubiquitous	*CXorf15*	
	CYorf15B	Chromosome Y open reading frame 15B	1	Ubiquitous	*CXorf15*	
	SMCY	SMC (mouse) homologue Y	1	Ubiquitous	*SMCX*	
	EIF1AY	Translation initiation factor 1AY	1	Ubiquitous	*EIF1AX*	
	RPS4Y2	Ribosomal protein S4Y isoform 2	1	Ubiquitous	*RPS4X*	
Ampliconic	*TSPY*	Testis-specific protein Y	~35	Testis		
	VCY	Variable charge Y	2	Testis	*VCX*	
	XKRY	XK related Y	2	Testis		
	CDY	Chromodomain Y	4	Testis		*CDYL*
	HSFY	Heat shock transcription factor Y	2	Testis		
	RBMY	RNA binding motif Y	6	Testis	*RBMX*	

MSY sequence class	Gene symbol	Gene name	Number of copies	Tissue expression	X-linked homologue	Autosomal homologue
	PRY	PTP-BL related Y	2	Testis		
	BPY2	Basic protein Y2	3	Testis		
	DAZ	Deleted in azoospermia	4	Testis		*DAZL*
TOTAL			~78			

Y chromosome structure in other mammals

Monotremes, the earliest diverging mammalian lineage, have heteromorphic, but very large, X and Y chromosomes, which pair over the entire short arm of the X and the long arm of the Y (Murtagh 1977). In contrast, the Y chromosome of marsupials (a lineage that diverged from eutherians later than monotremes) is very small, only about 12 Mb, containing a handful of genes. It lacks a PAR and does not recombine with the X. The X chromosome is also small in marsupials, about 3% of the haploid genome (Sharp 1982; Toder et al. 2000b).

In most eutherians, the Y chromosome is one of the smallest chromosomes in the genome. There are generally a small euchromatic part and a heterochromatic region that can vary in size between species (Leite-Silva et al. 2003; Marchal et al. 2003; Nova et al. 2002; Toder et al. 1997; Wolf et al. 1965). The euchromatic part contains few genes among them the *SRY* that is present in most eutherian mammals (Marshall Graves 2002a; Toder et al. 2000a; Waters et al. 2001) and marsupials (Foster et al. 1992). *DFFRY*, *SMCY*, *DBY*, *ZFY* and *UTY* are five other genes present in many mammalian orders (Murphy et al. 1999) (Paper I). They seem to have important male specific functions, since mutations in this region result in severe early blockage in spermatogenesis (Mazeyrat et al. 1998).

The gene content and the relative location of the genes on the Y chromosome vary between species. An example is the amelogenin gene that is located in the PAR of cattle, but outside the PAR in humans (Liu et al. 2002). In mice the amelogenin gene is only present as an X copy (Chapman et al. 1991) while in monotremes and marsupials, it is an autosomal gene (Watson et al. 1992).

The *UBE1* gene shows all stages of degeneration in different mammalian lineages from pseudoautosomal in monotremes to absence in the Old World monkeys (including humans) and marmoset. In all other mammals analyzed *UBE1* is a differentiated X and Y gene (Mitchell et al. 1998).

The number of gene copies can also vary between species. In mice, for example, two copies of *ZFY*, *Zfy1* and *Zfy2* have been found while humans have only one (Mazeyrat et al. 1998) and in the Microtidae family the numbers of *SRY* genes can vary from one to fifteen (Bullejos et al. 1997; Bullejos et al. 1999). Many human genes also exist in multiple copies, and in cattle, *Tspy*, is repeated over 1200 times (Matthews and Reed 1992; Vogel et al. 1997).

Y chromosomal rearrangements are quite common between and within species (Burgoyne et al. 1998; Cockwell et al. 2003; Glaser et al. 1998; Quilter et al. 2002). A pericentric inversion of the Y distinguishes the closely related species *Bos taurus* and *Bos indicus* where the position of the centromere in *Bos taurus* is metacentric/ submetacentric and in *Bos indicus* acrocentric (Potter and Upton 1979) (Figure 7)

In two species of the mole vole (*Ellobius*) the whole Y chromosome has been lost (Just et al. 2002; Just et al. 1995; Vogel et al. 1998). In the *E. lutescens* both sexes are X0 and in *E. tancrei* both sexes are XX with one X inactive. Obviously a new sex determining system took over in an ancestor of *Ellobius* since the male determining gene, *SRY*, is lost in these two species. *SRYs* controlling function in the sex-determining pathway is probably overtaken by another gene located on an autosome. Accumulation of new variants around this new sex-determining gene might have started the process of sex chromosome differentiation over again (Marshall Graves 2002c).

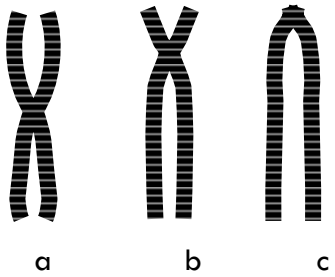

Figure 7*. The classification of chromosomes depend on the position of the centromere,* ***(a)*** *metacentric,* ***(b)*** *acrocentric or* ***(c)*** *telocentric.*

Below I will describe Y chromosome structure in one of my study species, the field vole.

Field vole *(Microtus agrestis)*

The field voles' sex chromosomes are extraordinarily large compared to the autosomes; the X chromosome is three times larger and the Y chromosome is twice the size of the largest autosome (Matthey 1949; Ohno 1967a). The size of the field vole sex chromosomes is apparent when compared to the human sex chromosomes (Figure 8). About three fourths of the field vole X chromosome constitutes of heterochromatin, accounting for 20% of the total female haploid genome (Wolf et al. 1965). The Y carries a similar amount of heterochromatin. The euchromatic portion of the X chromosome represents about 5% of the genome and corresponds to the "original" X, but the euchromatic region of the Y chromosome is, as in most mammals, very small. A striking feature of the field vole sex chromosomes is the absence of synapses, i. e. association between the X and Y in meiosis, a prerequisite for crossing over and recombination (Ashley et al. 1989).

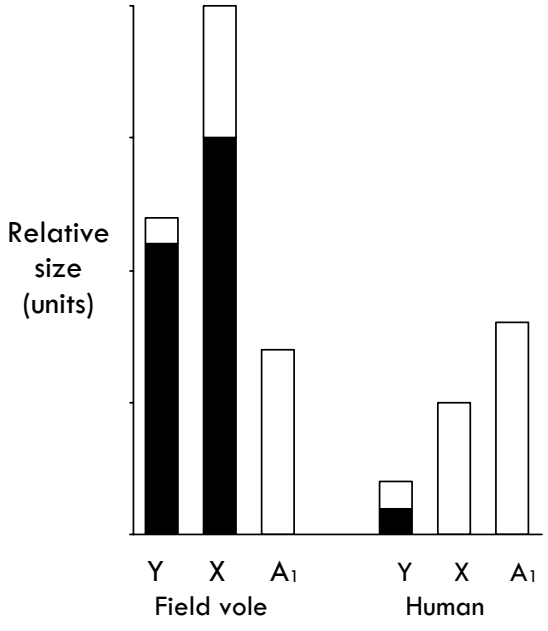

***Figure 8**. Mean size of chromosomes within chromosomal classes in field vole and humans. White is euchromatin and black is heterochromatin.*

An unusual finding in this species is the existence of two karyotypically different Y chromosomes. The standard Y is acrocentric but in southwestern Sweden, male field voles carry a subtelocentric Y. This Y chromosome, called the Lund Y, is the result of a pericentric inversion (Fredga and Jaarola 1997).

Y chromosome diversity

Intraspecific diversity as well as interspecific divergence can be estimated by pi (π) or theta (θ). Theta is the proportion of nucleotide sites that are polymorphic in a sample. Pi is the average number of nucleotide differences per site between two randomly chosen sequences in the sample. While θ is a measure of nucleotide polymorphism in a sample and can be corrected for by sequence length and sample size, π measures the nucleotide diversity with regard to frequencies of different alleles. Under neutrality, these estimates should be the same but selection and population structure will affect the estimates of π and θ in different ways (Li 1997).

Levels of genetic variation in the Y chromosome may differ from that of the rest of the genome for a number of reasons discussed below.

Effective population size

In sexual populations, half of the alleles are derived from females and half from males. The number of chromosome variants maintained in the population is, among other things, dependent on the effective population size (N_e) of each chromosome. In an ideal population the relationship between Y, X and autosomes is N_e:$3N_e$:$4N_e$ suggesting an expected 1:3:4 relationship in diversity between the different chromosomes.

Mating systems

Different mating systems can cause differences in effective population size. Skewed mating systems, for example in polygynous species, where one male mates with many females, will affect the relative difference in effective population size between chromosomes. For example; the Y:X:autosome relationship when one male mates with two females will be 1:5:6. If the ratio

of females to males is increased to ten (which is common in some species (McComb and Clutton-Brock 1994; Roed et al. 2002) the relationship would be 1:21:22.

Selection

New mutations can be neutral, advantageous or disadvantageous. The probability of fixation or elimination of the mutation in the population depends on the relative fitness of the new phenotype. Exceptions occur for balancing selection, overdominance (where heterozygotes are favored) and in limited populations. Negative selection will tend to eliminate disadvantageous mutants or genotypes from the population and is the prevailing type of selection since the majority of non-neutral mutations are deleterious or slightly deleterious. Positive selection increases the probability for an advantageous mutation to become fixed in the population (Li 1997). However, the chance of losing a new advantageous mutation from the population by random genetic drift (change in allele frequency due to chance) can still be high (Hartl and Clark 1997).

Selection at a locus will also affect linked sites. In the absence of recombination, selection will tend to reduce genetic variability at linked sites to the same extent as at the locus under selection. With recombination, the effect becomes gradually smaller as the rate of recombination between the selected locus and linked sites increases. In line with this thinking, levels of neutral variability have been shown to correlate with recombination rate in humans (Nachman 2001), mice (Nachman 1997), plants (Stephan and Langley 1998) and fruit flies (Begun and Aquadro 1992). The Y chromosome, which lacks recombination (except in the PAR), should be expected to have reduced variation as compared to recombining chromosomes. Selective sweeps and background selection may have severe affects on the MSY where all sites are linked compared to other genomic regions.

Selective sweeps

Selective sweep or the hitchhiking effect is an effect of positive selection where a favorable allele drives through the population to fixation together with its' linked loci. This will reduce the variation linked to the selected site and decrease the diversity in the population (Rice 1987). The impact of a selective sweep depends on the recombination rate and selection coefficient, the lower the recombination rate and/or the higher the selection coefficient, the larger is the genomic region affected by the sweep (Figure 9). In the Y chromosome where 95% of the sites are linked, an advantageous gene regulating a male specific trait, like one involved in spermatogenesis, may sweep through the population and eliminate all variation in the MSY (Roldan and Gomendio 1999; Wyckoff et al. 2000). Selective sweeps can bring about fixed Y chromosomes within a species and different between species, while mutations will only slowly produce new variants in a population. In contrast to the neutralists prediction of a positive correlation of intraspecific variation and interspecific divergence, positive selection can lead to uncoupling of levels of polymorphism and divergence (Li 1997).

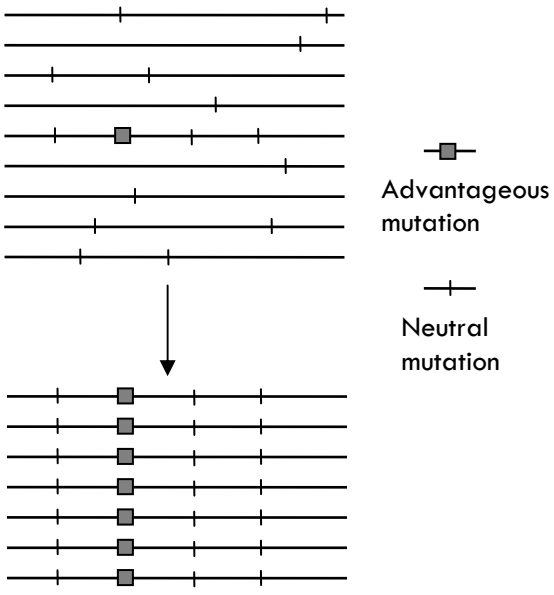

***Figure 9**. Selective sweep. An advantageous allele will drive through the population to fixation together with its linked sites.*

Background selection

Background selection is an effect of negative selection where deleterious mutations will be eliminated from the population together with their linked loci. This process, as with selective sweeps, will reduce variation in the region around the selected site (Figure 10). Similar to sweeps the impact of background selection depends on the recombination rate and the selection coefficient. Background selection is not thought to alter allele frequencies to the same extent as selective sweeps; indicating that the two types of selection can be distinguished from each other (Charlesworth et al. 1995). In a non-recombining region, mildly deleterious as well as weakly advantageous alleles will survive linked to each other. In the absence of a strongly advantageous mutation, a neutral or weakly selected mutation can only survive on a non-recombining chromosome (like Y) if there is no strongly deleterious mutation, otherwise it will be eliminated (Charlesworth 1994).

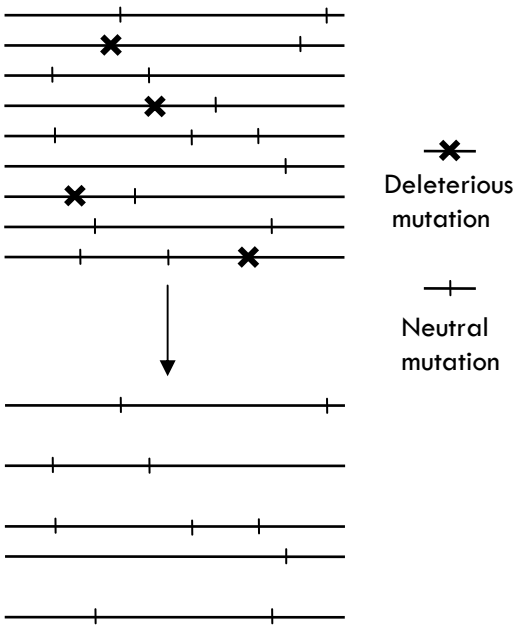

***Figure 10**. Background selection. Neutral mutations will disappear together with linked deleterious mutations when these are selected against.*

Sex specific mutation rates

Mutations are generated during DNA replication. The number of germ cell divisions differs between spermatogenesis and oogenesis. In oogenesis every mature oocyte has gone through a total of 24 cell divisions irrespective of the age of the female. In spermatogenesis however cell division is a continuous process, so the older the male the more cell divisions his sperms have gone through. For example in a 20 year-old man every sperm has gone through about 150 cell divisions and at the age of 40, 610 cell divisions (Hurst and Ellegren 1998).

The male to female mutation rate ratio, α_m, is mostly dependent on the skewed number of cell division in the germ lines of males and females. As α_m is generally larger than one, meaning that male germ cells mutate more frequently than female germ cells, more mutations in the Y chromosome than in other chromosomes can be predicted (Miyata et al. 1987). Estimates of α_m from X and Y comparisons suggest that it co-varies with the mean age of reproduction; rodent (α_m=2)(Chang et al. 1994) < felidae (α_m=4)(Pecon Slattery and O'Brien 1998) < primates (α_m=6)(Chang et al. 1996).

Other factors affecting Y diversity

Differences in migration between males and females can produce variation in the patterns of genetic differentiation detected in maternally and paternally inherited systems. In a patrilocal species, where female migrate more than males, this will imply less variation in the Y chromosome locally. In a global perspective, this will lead to higher differentiation in Y chromosome than in the maternally inherited mitochondrial DNA (mtDNA) (Seielstad et al. 1998).

Spermatogenesis and sperm mobility are energy demanding processes; therefore, the function of the mitochondria is vital for reproductive success. Deleterious mutations in the mtDNA that affect the energy production negatively would be expected to lead to impaired reproduction and, consequently, reduced effective population size among males. This will

lower the effective population size of Y chromosomes and reduce its diversity (Gemmell and Sin 2002).

The development of Y markers

Evolutionary or population genetic studies focused at male-specific patterns obviously require genetic markers from the Y chromosome. This may be in the form of polymorphic sequences for intraspecific studies, but may in principle concern any Y chromosome specific sequence.

In human and mouse the genome projects have been advanced producing a lot of Y chromosome sequence for these species (International Human Genome Sequencing Consortium 2002; Skaletsky et al. 2003; Waterston et al. 2002). In the absence of large-scale genome sequence information, alternative approaches are needed.

The two most popular methods are random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP). RAPD is a very fast method with low cost (Hadrys et al. 1992) but hard to reproduce since the conditions must be exactly the same from time to time. AFLP on the other hand is very consistent but time consuming and more expensive than RAPD (Mueller and Wolfenbarger 1999). Both methods can be used with an almost unlimited set of primers that will render an inexhaustible source of possibilities to develop Y markers.

Random amplified polymorphic DNA (RAPD)

Random amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR) uses a single short primer, which will anneal to various places in the genome and amplify fragments of different length in the PCR (Williams et al. 1990). These fragments can be separated on an agarose gel and male-specific bands can be isolated and sequenced to develop locus-specific Y-linked markers (Gutierrez-Adan et al. 1997; Olivier and Lust 1998; Wardell et al. 1993).

Amplified fragment length polymorphism (AFLP)

Double stranded adapters are ligated onto digested DNA. These adapters are complementary to the primers used in the next step, which is a preselective amplification step. The complementary primers have an extra base at the 3´ end (P-n), which extends into the genomic DNA, amplifying only one fourth of the fragments. After this, a second amplification step is performed with primers similar to the preselective one but with two additional bases at the 3´ end (P-nnn) amplifying only 1/16 of the fragments from the preselective PCR. The results can then be displayed on a denaturing acrylamide gel either by silver staining or by radioactive labeled primers (Vos et al. 1995).

Sex linked markers can be identified when comparing AFLP profiles from a number of males and females. These bands can then be cut from the gel and sequenced. Although much more time consuming than the RAPD technique, the AFLP produces about 10 times more visible bands/primer and thereby greater chance of detecting sex-specific fragments (Griffiths and Orr 1999).

Other techniques

There are other methods that can be used to get Y specific sequence, for example reduced representation shotgun (RRS) (Altshuler et al. 2000) together with flow sorting (Mullikin et al. 2000). This is an approach where the Y chromosome is isolated and digested with a restriction enzyme. The fragments are then separated on an agarose gel and bands appropriate for insertion and sequencing from clones are sliced out. This method will pick up sequence variation, indels and SNPs between individuals, if a number of individuals were pooled together in the digestion reaction.

Applications

There are two classes of chromosomal markers with respect to mutation rates, “unique” and recurrent event markers. The “unique” mutations (with a mutation rate about $1x10^{-9}$ per site per generation) are considered to occur only once in the evolutionary history of a species for example, SNPs and insertion-deletion events (Hammer 1994; Thomson et al. 2000). In the MSY these binary polymorphisms can be combined into haplogroups with a monophyletic origin, and the phylogenetic relationship can thus be depicted in a single most parsimonious tree (a tree with the least mutations to explain the topology).

Microsatellites or short tandem repeats (STRs) are tandemly repeated units of 1-6 bp scattered in the genome. They are considered to be recurrent event markers because of their high mutation rate (about $2x10^{-3}$)(Ellegren 2000; Kayser et al. 2000). Their high level of polymorphism makes STRs useful in, for example for paternity testing and to detect recent population differentiation. STRs on the Y chromosome (outside the PAR region) can be combined into haplotypes, but because of the high level of homoplasy (identity by state but not by descent) phylogenetic reconstruction is difficult. Given their different levels of polymorphism, combining information from the two classes of mutations makes the Y chromosome a powerful tool to detect male population structure and differentiation on different time-scales (Jobling and Tyler-Smith 1995; Mitchell and Hammer 1996). Unique and recurrent mutations can be combined to reveal Y chromosome genealogies. Unique mutations define deep lineages (haplogroups) while recurrent mutations define terminal haplotypes (de Knijff et al. 1997). In this way many recurrent events in a population causing loops in a network can be resolved if considered in their respective haplogroups because each haplogroup is founded by a single male with no variation at the multiallelic loci.

Anthropology

As the number of Y chromosome studies in mammals other than humans is limited, this chapter will focus on anthropological studies using the Y chromosome.

Patterns of migration

In global population studies, more geographic differentiation in the Y chromosome than in the mtDNA is generally detected (Kayser et al. 2000; Seielstad et al. 1994; Seielstad et al. 1998; Underhill et al. 1997) (Figure 11). To explain the higher differentiation on Y than mtDNA, difference in migration rate between males and females can be invoked. This is consistent with a patrilocal society where children are mostly brought up at the place where the father originates if the parents are from different places. To explain the result an eight times higher prehistoric migration rate of females than males was estimated. Over many generations this will be observed globally by genetically homogenous mtDNA but differentiate Y chromosomes (Seielstad et al. 1998).

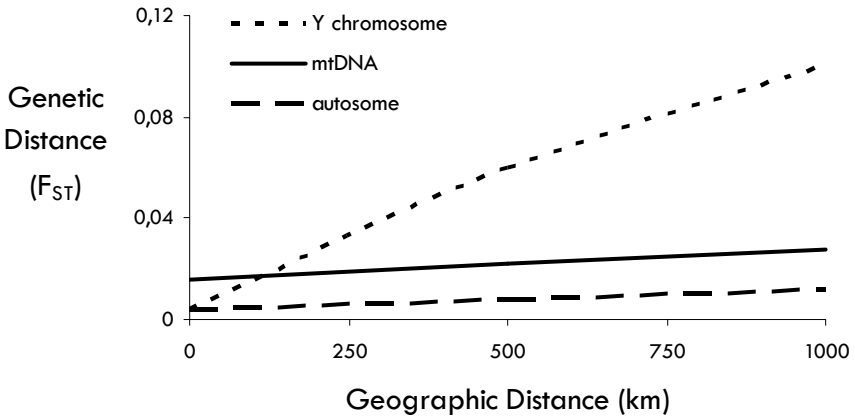

***Figure 11**. Relationship between genetic distance (F_{ST}) and geographic distance (km) in humans for Y chromosome, mtDNA and autosome. Modified from Seielstad et al (1998).*

Genetic structure in European populations

Genetic studies in European males show a strong correlation between Y haplogroups and geography but low correlation between Y haplogroups and language (Malaspina et al. 1998; Poloni et al. 1997; Quintana-Murci et al. 1999; Rosser et al. 2000).

Iceland was one of the last landmasses in Europe to be colonized by humans. According to the book of settlements (Pálsson H 1972), Norse Vikings settled down in 870-930 A.D. accompanied by some people from the British Isles. When studying maternal ancestry using mtDNA lineages, Helgason et al (Helgason et al. 2000a) found a closer relationship between Icelanders and populations from the British Isles than to the Scandinavian population. However paternally inherited Y chromosome markers suggested a main Norwegian origin with only about 20% contribution from the British Isles (Helgason et al. 2000b). These results could indicate that Norwegian Vikings (mostly men) took slaves or wives from their colonies in the British Isles (mostly women) when they colonized Iceland.

Surnames and Y lineages

Surnames are mainly inherited from fathers to children in almost all societies. This tradition is almost 5000 years old in China, 700 years in England, but only 68 years in Turkey (Jobling 2001). This social custom implies a relationship between the Y chromosome genealogies and male surnames (Figure 12a). The correlation between a specific surname and a Y chromosome lineage is of course dependent on there having been no illegitimate matings since this most certainly will introduce chromosomes from other surname groups (Figure 12b). Other reasons for uncoupling of surnames and Y chromosome haplotypes are that the surname does not have a unique origin like our Swedish –son names e.g Svensson or Andersson (Figure 12c) and that each establishment of a new surname has not started with a unique haplogroup (Figure 12d).

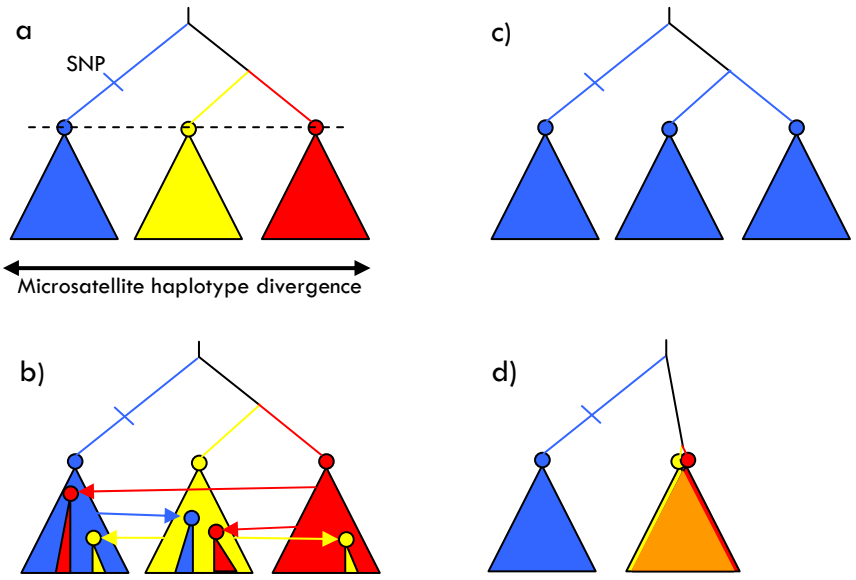

***Figure 12**. Relationship between Y-chromosomal haplotypes and three different surnames, represented by blue, yellow and red. Coloured circles represent founders and the cones represent microsatellite haplotype divergence. **a**) An ideal situation where each founders' haplotype are different from the others and blue founder is also differentiated by an SNP. The dashed line represent time of surname establishment and the arrow indicate microsatellite haplotype divergence. **b**) The effect of illegitimacy, surname adoption or maternally inherited surname on the correlation between surname and Y haplotypes **c**) Polyphyletic surname. **d**) The effects of close haplotype relationship where yellow and red haplotypes overlap and can not be distinguished from each other. Modified from Jobling (2001).*

The first example of linkage between a surname and a Y chromosome lineage is a French-Canadian family where the ancestor of 17 men with the same surname and the same rare Y chromosome translocation could be traced back to a Canadian immigrant in 1665. Interestingly, another family with a different surname shared the same Y chromosome translocation probably through illegitimacy around 1830 (Genest 1973).

The Jewish Cohanim priesthood is supposed to share patrilineal descent from Aaron who lived 3000 years ago. Y chromosome analyses within this group showed that the most common Y haplogroup, and its one-step microsatellite mutation neighbor, was present in more than 60% of the Cohanim priests but in less than 15% in the control group (Thomas et al. 1998).

The most famous paternity test is probably that of Thomas Jefferson, the third president of the United States, who shared a quite uncommon haplotype with males named Jefferson. The same Y chromosome haplotype is also present in the descendents of the children of Sally Hemings, Thomas Jeffersson's slave. (Foster et al. 1998)

Social structure

Hindu populations are divided into approximately 2000 castes grouped basically into five varna with different status. This structure decides social processes such as marriages, which are preferred between individuals from the same varna. However, a woman can marry a man from a higher varna but the opposite, that men marry women from a higher status, is strongly discouraged. The genetic effects of this hierarchic system can be seen comparing mtDNA and Y chromosome data. Genetic distance in mtDNA is consistent with "isolation by culture" where varnas with close status are genetically more similar than varnas from the highest and lowest status in concordance with women's limited but upward social mobility. In contrast, the Y chromosome genetic distance does not show any relationship with status following men's inability to move in the hierarchical system (Bamshad et al. 1998).

Medicine

Infertility

Already in 1976 the deletion of the long arm of the Y chromosome was known to be associated with spermatogenic failure (Tiepolo and Zuffardi 1976) but the critical regions were not described until 1996 (Vogt et al. 1996). Complete deletion of AZFa (azoospermia factor a) and AZFb in the proximal and middle Yq11 are both correlated with male infertility. The AZFa region includes for example *DBY*, *UTY, TMSB4Y* and *USP9Y* while the AZFb region includes *RBMY*, *PRY* and *TTY2*. However, partial deletions of these two regions and all types of deletions in AZFc in the distal part of

Yq11 show a variable phenotype from fertile to infertile, suggesting perhaps a compensation by X and/or autosomal homologues for these genotypes (Krausz et al. 2003). Although microdeletions in the Y chromosome are the most common cause of male infertility other factors such as variations in repeat sequences in multicopy genes, mutations and chromosomal rearrangements could all contribute to this phenotype.

Other mammals

The use of Y chromosome sequence analyses in genetic studies of mammals other than humans fall within a number of categories. Paternal gene flow has been analyzed in shrew (Lugon-Moulin and Hausser 2002) and wolves (Sundqvist et al. 2001) and interspecific Y chromosome variation was investigated in bovids (Edwards et al. 2000). The Y chromosome has also been used in phylogenetic studies in African buffalos (Van Hooft et al. 2002) and macaques (Tosi et al. 2003). In the field of conservation genetics Y chromosome analyses have shed light into hybridization of wolves and dogs (Vila et al. 2003) and the status of the Przewalski's horse (Wallner et al. 2003). Different demographic histories between humans and chimpanzees could be elucidated by analyzing Y chromosome diversities in these species (Stone et al. 2002). Below are examples of Y chromosome analyses that have been used to study the process of domestication.

Domestication

Domestic cattle are derived from aurochs, *Bos primigenius*, that diverged into two species *Bos taurus* and *Bos indicus* (Troy et al. 2001) from which our domestic cattle are descendants. Paternally and maternally inherited DNA suggest that most Asian domestic cattle breeds are hybrids between these two species and may also harbor traces of other bovine species existing in Asia today, like Bali cattle (*Bos javanicus*), gaur (*Bos gaurus*) and yak (*Bos grunniens*) (Kikkawa et al. 2003). The Indian breeds have experienced the least taurine influence especially Ongole and Sahiwal (Kumar et al. 2003).

It is thought that cattle were first introduced into Africa in the form of Hamitic Longhorns (*B. taurus*) arriving in the Nile Delta 6000 BC. A second introduction was by Shorthorns (*B. taurus*) around 2500 BC, before the introduction of *Bos indicus* (Epstein 1971). Although taurine cattle are most likely the oldest domesticate on the African continent, the indicine Y chromosome haplotype dominates the present population (70%) possibly because indicine cattle are better adapted to heat and drought (Hanotte et al. 2000). The authors suggest three main reasons that prevent the spread of the indicine allele throughout Africa. First the taurine adaptation to specific diseases like trypanosomosis carried by tsetse-infested fly, second preference by farmers for certain phenotypic traits and third geographic isolation (Hanotte et al. 2000). There has also been some evidence of an African centre of domestication from archeological (Grigson 1991) and mtDNA studies (Bradley et al. 1996; Troy et al. 2001).

Cattle were first introduced to America by Spanish conquerors in 1492. In a few years cattle spread all over the continent and nowadays almost all North and South American countries have Creole cattle, a native breed descendant of Iberian cattle (Giovambattista et al. 2000). These first cattle were of taurine origin, which thus dominates the male population in the South American countries. An exception is Brazil where zebu males were introduced in the eighteenth and nineteenth centuries, to improve native breeds in tropical regions (Giovambattista et al. 2000).

Summaries of Paper I-IV

In order to obtain Y chromosome markers useful across a range of mammalian species, I first developed conserved Y chromosome primers and tested them in 20 different species (Paper I). To study the degree of Y chromosome variability, I then contrasted levels of nucleotide diversity in gametologous X and Y chromosome sequences in five mammalian species (Paper II). The field vole and cattle were finally chosen to apply the Y chromosome markers, in studies of speciation and domestication respectively (Paper III and Paper IV).

Paper I

Y Chromosome Conserved Anchored Tagged Sequences (YCATS) for the Analysis of Mammalian Male-specific DNA

The amount of available Y chromosome sequence is scarce in most species. Most population genetic studies of natural populations of most species have been based on mtDNA or autosomal markers. In this study, we developed conserved Y chromosome markers based on comparative anchored tagged sequences (CATS) (Lyons et al. 1997). The CATS approach identifies conserved regions in exons from different species, in which primers then are designed to amplify across a range of species. In Paper I a modified approach called YCATS (Y chromosome conserved anchored tagged sequences) was used by aligning genes from X and Y cDNA sequences available in Genbank, mostly from human and mouse. We chose genes present in only one copy on Y (Lahn and Page 1997; Lahn and Page 1999a) to avoid multiple fragments in the PCR reaction. Primers were designed so that the base(s) at the 3' end differed between the X and Y sequences but

were similar within the respective chromosome. The primers were positioned in the exons flanking a short intron (in most cases less than 1000 bp in size) well suited for PCR.

Results

Forty-eight introns from six different Y chromosome genes were analyzed and screened in 20 mammalian species, covering 10 orders.

Two different PCR profiles were performed for each primer pair. A successful Y specific amplification was considered if male(s) showed a unique band that was not present in female(s).

Thirty-nine (of the 48) Y specific introns could be amplified in one or more species with a mean of 9.9±4.8 introns per species. On average 5.1±4.2 kb of sequence was obtained per species, which represents over 100 kilobases (kb) of mammalian Y chromosome DNA. Approximately 58 kb of the derived Y specific fragments were sequenced and compared with human X and Y sequences, which consistently confirmed Y chromosome origin. The success rate varied between different species being highest in chimpanzee (59 %) and field vole (41%) which are close relatives to the human and mouse whose sequences were used originally for primer design.

Discussion

Our success rate was unexpectedly low. Even for humans, which primers were designed for, had only 74% success and generally for the other species below 25%. The low rate in humans could be due to technical and chemical properties. The low success rate in the other species might be due to the attribute of the Y chromosome which may make cross-species amplification difficult. A male-biased mutation rate will obviously lead to higher synonymous substitution rates on Y-linked compared to autosomal or X-linked genes (Hurst and Ellegren 1998). A higher mutation rate could thus make cross-species amplification more difficult.

Another complicating factor is the dynamic evolution of the Y chromosome. Genes present in one species may be gone in others (Mitchell et al. 1998) or be duplicated (Mazeyrat et al. 1998). Apart from the species

where the Y chromosome is mapped, we do not know the genetic content of the Y chromosomes and failure to amplify a Y specific fragment may be due to circumstances like deletions or duplications of whole genes.

The Y chromosome genes used in this study have a gametologous copy on the X chromosome. These gene pairs (the X and Y copy) show different stages of differentiation. In order to avoid amplifying the X gametologue, genes from the older strata should be chosen. However, the differentiation of a non-recombining region is fast and therefore Y genes corresponding to stratum 1 might be very different between species making cross-amplification difficult. On the other hand, genes from stratum 4 might still recombine in some species and should therefore not be considered as templates for conserved Y markers.

Paper II

Low Levels of Nucleotide Diversity in Mammalian Y Chromosomes

Levels of nucleotide diversity are determined by factors such as selection, effective population size, rate of mutation and recombination. Sex chromosomes show distinct features in the relative magnitude of these evolutionary factors. The Y chromosome is subject to male-derived mutations only, is non-recombining and the effective population size is only one third of that of the X chromosome, which is present in both sexes and recombines in females.

Lower than expected intraspecific Y chromosome variation has been found in humans (Dorit et al. 1995; Shen et al. 2000; The International SNP Map Working Group 2001), Drosophila (McAllister and Charlesworth 1999) and plants (*Silene*) (Filatov et al. 2000). Is this a general trend also across other mammals and, if so, which factors would be responsible for this? To address these questions we screened for single nucleotide polymorphisms (SNP) on X and Y chromosome introns from populations of five different mammals: lynx, wolf, reindeer, cattle and field vole.

Results

We surveyed between 1.1-3.0 kb of X chromosome sequence and 0.7-3.5 kb of Y chromosome sequence per species for intraspecific variation. The number of segregating sites was lower on the Y chromosome than on the X for all five species and for three of them (lynx, reindeer and cattle) no intraspecific variation could be found on the Y chromosome. The other two species, the wolf and field vole, had two and four segregating sites, respectively, which corresponds to nucleotide diversities (π_Y) of $0.4x10^{-4}$ and $1.7x10^{-4}$. In the X chromosome, 1-7 segregating sites were found per species and the nucleotide diversity (π_X) ranged from $1.6x10^{-4}$ in lynx to $8.0x10^{-4}$ in field voles.

X and Y nucleotide diversity is not directly comparable because of differences in the effective population size and mutation rates. Since the effective population size is three times higher for X than Y, in a randomly mating population, the expected degree of variability should be three times higher on X than Y. However, this will be counteracted by male-biased mutation. Taking both effective population size and mutation rates into account, we still have lower than expected Y chromosome diversity estimates.

Discussion

Apart from the factors already mentioned, mating systems and mitochondrial effects are other mechanisms that lower male effective population size and thereby Y chromosome diversity.

To find out which factor is most important in shaping Y chromosome variability a comparison with birds is relevant. Birds have a reversed sex chromosome organization characterized by female heterogamety (female ZW and males ZZ). Studies on the female-specific W chromosome have revealed very low levels of intraspecific variability (Berlin and Ellegren 2001; Montell et al. 2001). It thus seems that the sex-limited chromosome (Y and W) generally experiences reduced levels of variability. Mating systems (high variance in male reproductive success) cannot explain low levels of variability in the female-specific W chromosome, nor can mutations in mtDNA. In birds, these factors would slightly lower the diversity on the Z

chromosome compared to autosomes and W due to differences in time spent in the male germ line. This points to selection as an important factor affecting genetic variation in the sex-limited chromosome (Ellegren 2003). Whether this is due to selective sweeps or background selection remains to be elucidated.

Paper III

Sex-linked Markers Propose a New Mammalian Species in Europe

The field vole, *Microtus agrestis*, is a common species with a continuous distribution range from Portugal in west Europe to Lake Baikal in Asia. Mitochondrial DNA studies of the cytochrome *b* gene have revealed three distinct phylogenetic lineages with largely allopatric distributions (Jaarola and Searle 2002; Jaarola and Searle 2004). These three populations, named western, eastern and southern after their geographical distribution, most likely reflect different glacial refugia and post-glacial recolonization routes. The eastern and western split is the most recent one and dates back to the last glacial period 50-83 kya (the last glacial period started 115 kya with a cold spell at 60-75 kya (Andersen and Borns 1997)). The western and eastern populations probably derive from refugial areas in the Carpathians and southern Urals, respectively. These two populations will here be referred to as the northern population. The southern lineage is more distantly related, with an approximate separation 0.6-1 Mya. The southern population has thus remained isolated from the northern population for several glacial periods; its present distribution probably reflects late glacial expansion from Iberian refugia.

I investigated paternally (Y chromosome) and biparentally (X chromosome) inherited sequences to address the phylogenetic relationships in the field vole. The object of this study was to investigate whether the high degree of differentiation in mtDNA was matched by nuclear divergence and whether the two major field vole lineages could be considered separate species.

Results

Altogether 27 male field voles were analyzed for 5.1-5.3 kb of Y chromosome sequence. In this dataset, 49 variable sites were found distributed on 39 SNPs and 10 indels. These polymorphisms defined 14 haplogroups. A highly polymorphic compound pentanucleotide microsatellite was also observed and the absence/presence of one of the motifs was used as an additional bi-allelic character in the maximum parsimony **(**MP) analyses.

For the X chromosome 2.7-2.8 kb sequence was surveyed in 25 female field voles. The thirty polymorphic sites revealed were divided into 26 SNPs and 4 indels, defining 14 haplogroups. Nucleotide diversity for the X and Y chromosome was estimated at 2.81×10^{-3} (π_X) and 1.90×10^{-3} (π_Y). These values are higher than published data for mouse, chimpanzee and human.

According to our phylogenetic analyses of the X and Y sequences, using both maximum parsimony (MP) and distance (neighbor-joining, NJ) methods, the field vole exhibits two highly divergent lineages. In fact, the majority of X and Y chromosome variation can be attributed to this split. The deep phylogenetic breaks in the X and Y chromosome genealogies display 100% bootstrap support and the divergence in substitutions and indels between lineages is significantly larger than variation within lineages (random regrouping 1000 times, $P < 0.0001$). The genetic structure between the two groups is also evident from only substitution data (Y $F_{ST} = 0.945$ ($P < 0.0001$) and X $F_{ST} = 0.894$, ($P < 0.0001$)).

The net divergence between the lineages, defined by the phylogenetic trees, estimated from intron sequences was 0.68 % for the X chromosome and 0.70 % for the Y chromosome. The nucleotide diversity within the two lineages was estimated at 0.16×10^{-3} (π_X) and 0.15×10^{-3} (π_Y) for the southern lineage and 0.79×10^{-3} (π_X) and 0.32×10^{-3} (π_Y) for the northern lineage.

To better map the distribution range of the two lineages and screen for possible hybrids, we analyzed an additional 36 individuals for two X chromosome markers and 28 males for one Y chromosome marker. The results show that both the X and Y chromosome lineages have distinct southern and northern geographical distributions. Thus, both systems recover the two major phylogroups previously described for mtDNA.

The individuals analyzed for one Y and two X markers (as well as cytochrome *b*) included 19 field voles (including 12 males) from a 50 km transect in the area of contact. These results, although preliminary, indicate that the contact zone is narrow and that little or no gene flow is occurring between the two field vole lineages.

Discussion

The slower evolving nuclear sequences from the X and Y chromosome confirm the deep phylogenetic split of the northern and southern lineages previously described for mtDNA (Jaarola and Searle 2002).

Thus, the field vole displays two highly divergent lineages genetically separated by reciprocal monophyly not only in maternally (mtDNA), but also paternally (Y chromosome) and biparentally (X chromosome) inherited systems. The fact that monophyly is also evident in the X chromosome, gives strong support for an ancient separation of the southern and northern field vole lineages. X-linked genes takes much longer time to reach reciprocal monophyly because, the effective population size is approximately three times larger than for Y or mtDNA linked genes (Avise 2000).

The striking similarity of the phylogeographic patterns for all genetic markers investigated, demonstrates that the southern and northern phylogroups in the field vole represent independent evolutionary trajectories.

All individuals analyzed in this study were found to be of pure northern or southern origin. With the three differently inherited markers, we would have discovered a first generation hybrid directly and past hybridization would have been detected as recombination in the X chromosome. We did, however, not find any signs of hybridization even in the area of contact.

According to the reciprocally monophyletic gene trees and the congruent geographic distributions for all markers investigated, the two lineages represent phylogenetically defined species. Furthermore, our preliminary analyses of the area of contact indicate post-reproductive isolation which according to Mayr (1942), defines biological species. We therefore suggest that the field vole, *Microtus agrestis*, should be reclassified as two different species with northern and southern distributions in Europe.

Paper IV

Early Cattle Husbandry:
Y Chromosome Haplogroups Indicate
Aurochs Introgression in Europe

Domestication of wild animals and plants is the most important event in human history and made it possible for people to settle down and develop sophisticated cultures (Clutton-Brock 1999; Giuffra et al. 2000; Loftus et al. 1994b; Sherratt 1997; Trut 1999; Wang et al. 1999; Vila et al. 2001). Cattle was a crucial species in these early farmers lifestyle providing meat, milk and motive power. It was domesticated from its wild progenitor the aurochs (*Bos primigenius*), which was once widespread in Europe, northern Africa and southern Asia (Clutton-Brock 1999).

The archeological evidence available suggest that *Bos taurus* was domesticated in the Eastern Mediterranean region about 10 000 years ago. This is in agreement with genetic data. Maternally inherited mtDNA from ancient as well as modern cattle breeds suggest at least two independent domestication events, *Bos taurus* in Anatolia or the Middle East and *Bos Indicus* in Asia. An independent domestication in north of Africa has also been suggested (Bradley et al. 1996; Loftus et al. 1994a; Troy et al. 2001).

No evidence of independent local domestication or introgression from aurochs has been detected in mtDNA studies of ancient cattle and aurochs (Bailey et al. 1996). However, the morphological similarities between some Neolithic specimens of *taurus* and *primigenius* in central Europe and the occurrence of both species at archeological sites may suggest that hybridization could occur. Our knowledge of Neolithic cattle keeping also allows for this to happen. Mitochondrial data would, however, not be suitable to trace the process, since aurochs bull introgression would not be visible in maternally inherited DNA. In order to address the question of hybridization, Y chromosome variation was screened for in modern breeds and the variants were tested in ancient material from Europe.

Results

In 3.5 kb of intronic Y chromosome sequence from 20 European bulls, two haplogroups were found separated by a two bp indel and a transversion (A/C). The two haplogroups Y1 and Y3 were screened in 149 bulls from 42 different European breeds. A north-south distribution of the haplogroups was detected with the Y1 haplogroup more common in breeds from northern Europe and the Y3 haplogroup mostly in breeds from southern Europe.

In the same 3.5 kb of Y chromosome sequence, screened in a bull from full-bred Sahiwal (*Bos indicus*), four transitions and a microsatellite length polymorphism were found to distinguish between *indicus* and *taurus* haplogroup. All the variable sites found in the 3.5 kb Y chromosome sequence were analyzed in 46 African cattle from six breeds. The three different haplogroups, *indicus*, *taurus* Y1 and *taurus* Y3, were found with the same distribution of indicine and taurine haplogroups as in Hanotte et al. (2000).

Thirty-four ancient Bos samples dated between 5000-10000 BP and originating from Italy, Austria, Germany and Sweden was analyzed for the two variable sites found in modern *taurus* and for two sites distinguishing *indicus* and *taurus*. Not all sites could be typed for all individuals but considering the age of the material and that it is nuclear DNA, from a haploid chromosome, the success rate (21) was very good. Out of the successfully typed samples, five proved to be domesticated Neolithic cattle. Among the sixteen aurochs samples that were analyzed for at least one site, fifteen showed haplogroup Y1 or Y2 (an intermediate between Y1 and Y3 could be typed with certainty in one case) and one was consistent with Y3 or an intermediate between the *indicus* haplogroup and Y3.

Discussions

Mitochondrial data suggest only a limited number of domestication events, including the domestication of *Bos indicus* and *Bos taurus*. The split between these two lineages is estimated from mtDNA to be well above 100000 years ago, long before domestication.

The polymorphic sites separating the Y1 and Y3 type in *Bos taurus* are both slow evolving changes (transversion and indel) compared to the

variable sites distinguished between *Bos indicus* and *Bos taurus* (transitions and microsatellite). Thus, it is not possible to say which split is the older, the one between Y1 and Y3 or the one between Y3 and the *indicus* haplogroup, only from the polymorphic sites. All three variants were most certainly present in the wild population suggesting at least three independent domestications of bulls or hybridization between domesticates and aurochs bulls.

Early husbandry of cattle included free roaming herds probably with someone guarding them instead of fences. This loose husbandry would allow wild aurochs males to reproduce with the domesticated females. The geographical distribution of Y1 in modern cattle and its presence in ancient European aurochs suggest a high frequency of local hybridization between domesticated cows and wild aurochs bulls in the centre of Europe. This is a likely scenario, since both farmers and aurochs populated the central of Europe. The slightly bigger aurochs bull may have reproduced with domestic cows leaving DNA print to the next generation, while an aurochs cow would have to be domesticated in order to leave a DNA print to modern breeds.

The presence of Y1 and Y3 type in modern African cattle is complicated to interpret, it could be the result of trading domestic animals in the last centuries. Domestic cattle were crucial for the Cape colony, and farmers from northern Europe (what became the Boer community) may have used African as well as European cattle to support their life stock. To elucidate the African cattle history material from indigenous pre-colonial African bulls are needed.

Our material of aurochs was limited to Europe but introgression could have occurred where both *primigenius* and *taurus*/*indicus* were present at the same time.

The Future

I believe that Y chromosome sequence analyses have just begun in mammals. The Y chromosome can be used to address many questions such as; evolution of a male specific chromosome, paternal migration, phylogenetic reconstruction, paternity and molecular evolution in a haploid chromosome.

More specifically this thesis has generated two additional Y chromosome studies concerning *Microtus* and *Bos Taurus* as described below.

1. I would analyze the Y chromosome genes in the Microtidae family and look for functional and degenerated haplotypes. In two *Ellobius* species and one *Tokudaia* species, sex is determined without a Y chromosome or the SRY gene. In the Microtidae family the number of SRY gene copies varies, from one functional male-specific copy in the *M. agrestis* up to 15 different sequences present in both X and Y chromosome in the *M.cabrerae* (Bullejos et al. 1997; Bullejos et al. 1999). My hypothesis is that since SRY is the male determining gene and therefore should degenerate last of the genes on the Y chromosome, it would be interesting to correlate the number of SRY copies and functionality to other Y specific genes, in order to find evidence for this hypothesis.
2. It would also be useful to type more European cattle, for the Y differences, in order to confirm the north and south gradient in bulls detected in Paper IV. In domestic animals with strict breeding programs, like cattle, the genetic variation can detect evidence of origin, breeding strategy, domestication etc. It would be interesting to screen more cattle breeds from Europe and the rest of the world to identify the origin and distribution of these two male lineages. A larger sample would be very valuable in

order to study the degree of genetic exchange between breeds. An example of the possible effect of male-biased breeding strategies in domestic animals is clearly expressed by thoroughbred horses, where pedigree analyses indicate that one single stallion is responsible for 95% of all paternal lineages (Lindgren et al.). The study of Y chromosome diversity in cattle will show the impact of sex-biased breeding in this domestic species.

Acknowledgements

Thanks to:
Mattias, my love and companion, for being the best father to our wonderful daughter Eleonor, for your love for me, your cooking and writing useful computer programs which have made many analyses a lot easier.

Hans Ellegren, my professor/supervisor, for giving me the opportunity to be a PhD student, introducing me to very interesting Y chromosome studies and for letting me do all the side projects that I come up with, not always ending with a great result.

Mum and dad, my loving parents, for all your support, not only financial and for letting me explore my curiosity and "uppfinningsrikedom" in the kitchen when I was smaller, inventing new recipes, not always edible.

Hannah Sundström and Jesper Brohede, my office mates, for your encouragement, joyful company and very useful discussions in all areas. Especially thanks to Hannah, you always had a minute to discuss or help me with whatever I come up with.

Chris Walker, my first supervisor, for always taking time to talk and who taught me population genetics with great patience.

Carles Vila whose energy and knowledge we all admire and are grateful that you share with us whenever we need.

All my collaborators, especially Göran Spong, Maarit Jaarola, Anders Götherström and Cia Anderung for sharing material with me, your interest in the results and enthusiastic discussions.

Lori-Jayne Lawson Handley, Jennifer Leonard and Hannah Sundström for reading and commenting on the thesis.

The conservation group, Anna-karin, Frank, Jennifer, Eva, Karin, Susanne, Carles, Cia, Annika and temporary students, for listening patiently to me, your interesting suggestions and always educational discussions on various subjects, especially my weird results and projects.

And finally thanks to everyone in the department of evolutionary biology for being such a nice team.

References

Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, and Lander ES. 2000. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* **407:** 513-6.

Andersen, and Borns. 1997. *The ice age world*. Scandinavian University press, Oslo.

Ashley T, Jaarola M, and Fredga K. 1989. The behavior during pachynema of a normal and an inverted Y chromosome in Microtus agrestis. *Hereditas* **111:** 281-94.

Avise JC. 2000. *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge.

Bailey JF, Richards MB, Macaulay VA, Colson IB, James IT, Bradley DG, Hedges RE, and Sykes BC. 1996. Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species. *Proc R Soc Lond B Biol Sci* **263:** 1467-73.

Bamshad MJ, Watkins WS, Dixon ME, Jorde LB, Rao BB, Naidu JM, Prasad BV, Rasanayagam A, and Hammer MF. 1998. Female gene flow stratifies Hindu castes. *Nature* **395:** 651-2.

Barton NH, and Charlesworth B. 1998. Why sex and recombination? *Science* **281:** 1986-90.

Begun DJ, and Aquadro CF. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356:** 519-520.

Berlin S, and Ellegren H. 2001. Evolutionary genetics. Clonal inheritance of avian mitochondrial DNA. *Nature* **413:** 37-8.

Bradley DG, MacHugh DE, Cunningham P, and Loftus RT. 1996. Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci U S A* **93:** 5131-5.

Bullejos M, Sanchez A, Burgos M, Hera C, Jimenez R, and Diaz de la Guardia R. 1997. Multiple, polymorphic copies of SRY in both males and females of the vole Microtus cabrerae. *Cytogenet Cell Genet* **79:** 167-71.

Bullejos M, Sanchez A, Burgos M, Jimenez R, and Diaz De La Guardia R. 1999. Multiple mono- and polymorphic Y-linked copies of the SRY HMG-box in microtidae. *Cytogenet Cell Genet* **86:** 46-50.

Burgoyne PS, Mahadevaiah SK, Perry J, Palmer SJ, and Ashworth A. 1998. The Y* rearrangement in mice: new insights into a perplexing PAR. *Cytogenet Cell Genet* **80:** 37-40.
Chang BH, Shimmin LC, Shyue SK, Hewett-Emmett D, and Li WH. 1994. Weak male-driven molecular evolution in rodents. *Proc Natl Acad Sci U S A* **91:** 827-831.
Chang BH-J, Hewett-Emmett D, and Li W-H. 1996. Male-to-female ratios of mutation rate in higher primates estimated from intron sequences. *Zoological Studies* **35:** 36-48.
Chapman VM, Keitz BT, Disteche CM, Lau EC, and Snead ML. 1991. Linkage of amelogenin (Amel) to the distal portion of the mouse X chromosome. *Genomics* **10:** 23-8.
Charlesworth B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet Res* **63:** 213-27.
Charlesworth B, and Charlesworth D. 2000. The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci* **355:** 1563-72.
Charlesworth D, Charlesworth B, and Morgan MT. 1995. The pattern of neutral molecular variation under the background selection model. *Genetics* **141:** 1619-32.
Clutton-Brock J. 1999. *Natural History of Domesticated Mammals(2nd ed)*. Cambridge University Press, United Kingdom.
Cockwell AE, Jacobs PA, Beal SJ, and Crolla JA. 2003. A study of cryptic terminal chromosome rearrangements in recurrent miscarriage couples detects unsuspected acrocentric pericentromeric abnormalities. *Hum Genet* **112:** 298-302.
Crow JF. 1994. Advantages of sexual reproduction. *Dev Genet* **15:** 205-13.
Crow JF, and Kimura M. 1965. Evolution in sexual and asexual populations. *American Naturalist* **99:** 439-450.
de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Roewer L, and et al. 1997. Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med* **110:** 134-49.
Dorit RL, Akashi H, and Gilbert W. 1995. Absence of polymorphism at the ZFY locus on the human Y chromosome. *Science* **268:** 1183-5.
Edwards CJ, Gaillard C, Bradley DG, and MacHugh DE. 2000. Y-specific microsatellite polymorphisms in a range of bovid species. *Anim Genet* **31:** 127-30.
Ellegren H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet* **16:** 551-8.
Ellegren H. 2003. Levels of polymorphism on the sex-limited chromosome: a clue to Y from W? *Bioessays* **25:** 163-167.

Epstein H. 1971. *The Origin of the Domestic Animals of Africa*. Africana Publishing Corporation, New York.
Fagan EA. 1979. The gingham dog and the calico cat. *Nurs Adm Q* **3:** 57-62.
Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* **78:** 737-56.
Filatov DA, Moneger F, Negrutiu I, and Charlesworth D. 2000. Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature* **404:** 388-90.
Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R, Zerjal T, and Tyler-Smith C. 1998. Jefferson fathered slave's last child. *Nature* **396:** 27-8.
Foster JW, Brennan FE, Hampikian GK, Goodfellow PN, Sinclair AH, Lovell-Badge R, Selwood L, Renfree MB, Cooper DW, and Graves JA. 1992. Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. *Nature* **359:** 531-3.
Fredga K, and Jaarola M. 1997. The origin and distribution of the Lund Y chromosome in Microtus agrestis (Rodentia, Mammalia). *Hereditas* **126:** 25-34.
Gemmell NJ, and Sin FY. 2002. Mitochondrial mutations may drive Y chromosome evolution. *Bioessays* **24:** 275-9.
Genest P. 1973. Transmission hereditaire, depuis 300 ans, d'un chromosome Y á satellites dans une lignee familiale. *Annual Genetic* **67:** 72-85.
Giovambattista G, Ripoli MV, De Luca JC, Mirol PM, Liron JP, and Dulout FN. 2000. Male-mediated introgression of Bos indicus genes into Argentine and Bolivian Creole cattle breeds. *Anim Genet* **31:** 302-5.
Giuffra E, Kijas JM, Amarger V, Carlborg O, Jeon JT, and Andersson L. 2000. The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* **154:** 1785-91.
Glaser B, Grutzner F, Willmann U, Stanyon R, Arnold N, Taylor K, Rietschel W, Zeitler S, Toder R, and Schempp W. 1998. Simian Y chromosomes: species-specific rearrangements of DAZ, RBM, and TSPY versus contiguity of PAR and SRY. *Mamm Genome* **9:** 226-31.
Graves JA. 1995. The origin and function of the mammalian Y chromosome and Y-borne genes--an evolving understanding. *Bioessays* **17:** 311-20.
Griffiths R, and Orr K. 1999. The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. *Mol Ecol* **8:** 671-4.
Grigson C. 1991. An African origin for African cattle? - some archeological evidence. *African Archeological review* **9:** 119-144.
Gutierrez-Adan A, Cushwa WT, Anderson GB, and Medrano JF. 1997. Ovine-specific Y-chromosome RAPD-SCAR marker for embryo sexing. *Anim Genet* **28:** 135-8.

Hadrys H, Balick M, and Schierwater B. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Mol Ecol* **1:** 55-63.

Hammer MF. 1994. A recent insertion of an alu element on the Y chromosome is a useful marker for human population studies. *Mol Biol Evol* **11:** 749-61.

Hanotte O, Tawah CL, Bradley DG, Okomo M, Verjee Y, Ochieng J, and Rege JE. 2000. Geographic distribution and frequency of a taurine Bos taurus and an indicine Bos indicus Y specific allele amongst sub-saharan African cattle breeds. *Mol Ecol* **9:** 387-96.

Hartl DL, and Clark AG. 1997. *Principles of population genetics*. Sinauer Associates, Sunderland, Mass.

Hayashida H, Kuma K, and Miyata T. 1992. Interchromosomal gene conversion as a possible mechanism for explaining divergence patterns of ZFY-related genes. *J Mol Evol* **35:** 181-3.

Helgason A, Sigureth ardottir S, Gulcher JR, Ward R, and Stefansson K. 2000a. mtDNA and the origin of the Icelanders: deciphering signals of recent population history. *Am J Hum Genet* **66:** 999-1016.

Helgason A, Sigureth ardottir S, Nicholson J, Sykes B, Hill EW, Bradley DG, Bosnes V, Gulcher JR, Ward R, and Stefansson K. 2000b. Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet* **67:** 697-717.

Hurst LD, and Ellegren H. 1998. Sex biases in the mutation rate. *Trends Genet* **14:** 446-452.

International Human Genome Sequencing Consortium. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420:** 520-62.

Jaarola M, and Searle JB. 2002. Phylogeography of field voles (Microtus agrestis) in Eurasia inferred from mitochondrial DNA sequences. *Mol Ecol* **11:** 2613-21.

Jaarola M, and Searle JB. 2004. A highly divergent mitochondrial DNA lineage of Microtus agrestis in southern Europe. *Heredity* **92:** 228-234.

Jacobs PA, and Strongs JA. 1959. A case of human intersexuality having a possible XXY sex determining machanism. *Nature* **183:** 302-303.

Jobling MA. 2001. In the name of the father: surnames and genetics. *Trends Genet* **17:** 353-7.

Jobling MA, and Tyler-Smith C. 1995. Fathers and sons: the Y chromosome and human evolution. *Trends Genet* **11:** 449-56.

Just W, Baumstark A, Hameister H, Schreiner B, Reisert I, Hakhverdyan M, and Vogel W. 2002. The sex determination in Ellobius lutescens remains bizarre. *Cytogenet Genome Res* **96:** 146-53.

Just W, Rau W, Vogel W, Akhverdian M, Fredga K, Graves JA, and Lyapunova E. 1995. Absence of Sry in species of the vole Ellobius. *Nat Genet* **11:** 117-8.

Kalthoff K. 1996. *Analysis of biological development*. McGraw-Hill, New York.

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Kruger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, and Sajantila A. 2000. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am J Hum Genet* **66:** 1580-8.

Kikkawa Y, Takada T, Sutopo, Nomura K, Namikawa T, Yonekawa H, and Amano T. 2003. Phylogenies using mtDNA and SRY provide evidence for male-mediated introgression in Asian domestic cattle. *Anim Genet* **34:** 96-101.

Koopman P, Ashworth A, and Lovell-Badge R. 1991. The ZFY gene family in humans and mice. *Trends Genet* **7:** 132-6.

Krausz C, Forti G, and McElreavey K. 2003. The Y chromosome and male fertility and infertility. *Int J Androl* **26:** 70-5.

Kumar P, Freeman AR, Loftus RT, Gaillard C, Fuller DQ, and Bradley DG. 2003. Admixture analysis of South Asian cattle. *Heredity* **91:** 43-50.

Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Silber S, Oates R, Rozen S, and Page DC. 2001. The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat Genet* **29:** 279-86.

Lahn BT, and Page DC. 1997. Functional coherence of the human Y chromosome. *Science* **278:** 675-80.

Lahn BT, and Page DC. 1999a. Four evolutionary strata on the human X chromosome. *Science* **286:** 964-967.

Lahn BT, and Page DC. 1999b. Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome. *Nat Genet* **21:** 429-33.

Leite-Silva C, Santos N, Fagundes V, Yonenaga-Yassuda Y, and de Souza MJ. 2003. Karyotypic characterization of the bat species Molossus ater, M. molossus and Molossops planirostris (Chiroptera, Molossidae) using FISH and banding techniques. *Hereditas* **138:** 94-100.

Li W-H. 1997. *Molecular evolution*. Sinauer Associates Inc, Sunderland, Mass.

Lindgren G, Backström N, Swinburne J, Hellborg L, Einarsson A, Sandberg K, Cothran G, Vilà C, Binns M, and H. E. (In Press) Limited number of patrilines in horse domestication. *Nature Genetics*.

Liu WS, Mariani P, Beattie CW, Alexander LJ, and Ponce De Leon FA. 2002. A radiation hybrid map for the bovine Y Chromosome. *Mamm Genome* **13:** 320-6.

Loftus RT, MacHugh DE, Bradley DG, Sharp PM, and Cunningham P. 1994a. Evidence for two independent domestications of cattle. *Proc Natl Acad Sci U S A* **91:** 2757-61.

Loftus RT, MacHugh DE, Bradley DG, Sharp PM, and Cunningham P. 1994b. Evidence for two independent domestications of cattle. *Proc Natl Acad Sci U S A* **91:** 2757-61.

Lugon-Moulin N, and Hausser J. 2002. Phylogeographical structure, postglacial recolonization and barriers to gene flow in the distinctive Valais chromosome race of the common shrew (Sorex araneus). *Mol Ecol* **11:** 785-94.

Lyons LA, Laughlin TF, Copeland NG, Jenkins NA, Womack JE, and O'Brien SJ. 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nat Genet* **15:** 47-56.

Malaspina P, Cruciani F, Ciminelli BM, Terrenato L, Santolamazza P, Alonso A, Banyko J, Brdicka R, Garcia O, Gaudiano C, Guanti G, Kidd KK, Lavinha J, Avila M, Mandich P, Moral P, Qamar R, Mehdi SQ, Ragusa A, Stefanescu G, Caraghin M, Tyler-Smith C, Scozzari R, and Novelletto A. 1998. Network analyses of Y-chromosomal types in Europe, northern Africa, and western Asia reveal specific patterns of geographic distribution. *Am J Hum Genet* **63:** 847-60.

Marchal JA, Acosta MJ, Bullejos M, Diaz de la Guardia R, and Sanchez A. 2003. Sex chromosomes, sex determination, and sex-linked sequences in Microtidae. *Cytogenet Genome Res* **101:** 266-73.

Marshall Graves JA. 2002a. The rise and fall of SRY. *Trends Genet* **18:** 259-64.

Marshall Graves JA. 2002b. Sex chromosomes and sex determination in weird mammals. *Cytogenet Genome Res* **96:** 161-8.

Marshall Graves JA. 2002c. Sex chromosomes and sex determination in weird mammals. *Cytogenet Genome Res* **96:** 161-168.

Marshall Graves JA, and Shetty S. 2001. Sex from W to Z: evolution of vertebrate sex chromosomes and sex determining genes. *J Exp Zool* **290:** 449-62.

Matthews ME, and Reed KC. 1992. Sequences from a family of bovine Y-chromosomal repeats. *Genomics* **13:** 1267-1273.

Matthey R. 1949. Chromosomes sexuels geant chez un Campagnol, Microtus agrestis L. *Experientia* **5:** 72-73.

Maynard-Smith J. 1978. *The evolution of sex*. Cambridge University Press, Cambridge.

Mayr E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York.

Mazeyrat S, Saut N, Sargent CA, Grimmond S, Longepied G, Ehrmann IE, Ellis PS, Greenfield A, Affara NA, and Mitchell MJ. 1998. The mouse Y chromosome interval necessary for spermatogonial proliferation is gene dense with syntenic homology to the human AZFa region. *Hum Mol Genet* **7:** 1713-24.

McAllister BF, and Charlesworth B. 1999. Reduced sequence variability on the Neo-Y chromosome of *Drosophila americana americana*. *Genetics* **153:** 221-233.
McComb K, and Clutton-Brock T. 1994. Is mate choice copying or aggregation responsible for skewed distributions of females on leks? *Proc R Soc Lond B Biol Sci* **255:** 13-9.
Mitchell MJ, Wilcox SA, Watson JM, Lerner JL, Woods DR, Scheffler J, Hearn JP, Bishop CE, and Graves JA. 1998. The origin and loss of the ubiquitin activating enzyme gene on the mammalian Y chromosome. *Hum Mol Genet* **7:** 429-34.
Mitchell RJ, and Hammer MF. 1996. Human evolution and the Y chromosome. *Curr Opin Genet Dev* **6:** 737-42.
Miyata T, Hayashida H, Kuma K, Mitsuyasu K, and Yasunaga T. 1987. Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harb Symp Quant Biol* **52:** 863-867.
Montell H, Fridolfsson AK, and Ellegren H. 2001. Contrasting levels of nucleotide diversity on the avian Z and W sex chromosomes. *Mol Biol Evol* **18:** 2010-6.
Mueller UG, and Wolfenbarger LL. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* **14:** 389-394.
Muller HJ. 1964. The Relation of Recombination to Mutational Advance. *Mutat Res* **106:** 2-9.
Mullikin JC, Hunt SE, Cole CG, Mortimore BJ, Rice CM, Burton J, Matthews LH, Pavitt R, Plumb RW, Sims SK, Ainscough RM, Attwood J, Bailey JM, Barlow K, Bruskiewich RM, Butcher PN, Carter NP, Chen Y, Clee CM, Coggill PC, Davies J, Davies RM, Dawson E, Francis MD, Joy AA, Lamble RG, Langford CF, Macarthy J, Mall V, Moreland A, Overton-Larty EK, Ross MT, Smith LC, Steward CA, Sulston JE, Tinsley EJ, Turney KJ, Willey DL, Wilson GD, McMurray AA, Dunham I, Rogers J, and Bentley DR. 2000. An SNP map of human chromosome 22. *Nature* **407:** 516-520.
Murphy WJ, Sun S, Chen ZQ, Pecon-Slattery J, and O'Brien SJ. 1999. Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. *Genome Res* **9:** 1223-30.
Murtagh CE. 1977. A unique cytogenetic system in monotremes. *Chromosoma* **65:** 37-57.
Nachman MW. 1997. Patterns of DNA variability at X-linked loci in *Mus domesticus*. *Genetics* **147:** 1303-1316.
Nachman MW. 2001. Single nucleotide polymorphisms and recombination rate in humans. *Trends Genet* **17:** 481-485.
Nova P, Reutter BA, Rabova M, and Zima J. 2002. Sex-chromosome heterochromatin variation in the wood mouse, Apodemus sylvaticus. *Cytogenet Genome Res* **96:** 186-90.

Ohno S. 1967a. *Sex chromosomes and sex linked genes*. Springer, Berlin.
Ohno S. 1967b. *Sex chromosomes and sex-linked genes*. Springer, Berlin New York.
Olivier M, and Lust G. 1998. Two DNA sequences specific for the canine Y chromosome. *Anim Genet* **29:** 146-9.
Pálsson H EP. 1972. *The book of settlements: landnámabók*. University of Manitoba, Winnipeg.
Pecon Slattery J, and O'Brien SJ. 1998. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. *Genetics* **148:** 1245-55.
Pecon Slattery J, Sanner-Wachter L, and O'Brien SJ. 2000. Novel gene conversion between X-Y homologues located in the nonrecombining region of the Y chromosome in Felidae (Mammalia). *Proc Natl Acad Sci U S A* **97:** 5307-12.
Poloni ES, Semino O, Passarino G, Santachiara-Benerecetti AS, Dupanloup I, Langaney A, and Excoffier L. 1997. Human genetic affinities for Y-chromosome P49a,f/TaqI haplotypes show strong correspondence with linguistics. *Am J Hum Genet* **61:** 1015-35.
Potter WL, and Upton PC. 1979. Y chromosome morphology of cattle. *Aust Vet J* **55:** 539-41.
Quilter CR, Blott SC, Mileham AJ, Affara NA, Sargent CA, and Griffin DK. 2002. A mapping and evolutionary study of porcine sex chromosome genes. *Mamm Genome* **13:** 588-94.
Quintana-Murci L, Semino O, Minch E, Passarimo G, Brega A, and Santachiara-Benerecetti AS. 1999. Further characteristics of proto-European y chromosomes. *Eur J Hum Genet* **7:** 603-8.
Rice WR. 1987. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116:** 161-7.
Roed KH, Holand O, Smith ME, Gjostein H, Kumpula J, and Nieminen M. 2002. Reproductive success in reindeer males in a herd with varying sex ratio. *Mol Ecol* **11:** 1239-43.
Roldan ER, and Gomendio M. 1999. The Y chromosome as a battle ground for sexual selection. *Trends in Ecology and Evolution* **14:** 58-62.
Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C,

and Jobling MA. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* **67:** 1526-43.
Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, and Page DC. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* **423:** 873-6.
Sandstedt SA, and Tucker PK. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. *Genome Res* **14:** 267-72.
Seielstad MT, Hebert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, and Cavalli-Sforza LL. 1994. Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum Mol Genet* **3:** 2159-61.
Seielstad MT, Minch E, and Cavalli-Sforza LL. 1998. Genetic evidence for a higher female migration rate in humans. *Nat Genet* **20:** 278-80.
Sharp P. 1982. Sex chromosome pairing during male meiosis in marsupials. *Chromosoma* **86:** 27-47.
Shen P, Wang F, Underhill PA, Franco C, Yang WH, Roxas A, Sung R, Lin AA, Hyman RW, Vollrath D, Davis RW, Cavalli-Sforza LL, and Oefner PJ. 2000. Population genetic implications from sequence variation in four Y chromosome genes. *Proc Natl Acad Sci U S A* **97:** 7354-9.
Sherratt, A. 1997 *Economy and society in prehistoric Europe.* Edinburgh University Press. Edinburgh.
Shetty S, Griffin DK, and Graves JA. 1999. Comparative painting reveals strong chromosome homology over 80 million years of bird evolution. *Chromosome Res* **7:** 289-95.
Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, and Page DC. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423:** 825-37.
Stephan W, and Langley CH. 1998. DNA polymorphism in lycopersicon and crossing-over per physical length. *Genetics* **150:** 1585-1593.
Stone AC, Griffiths RC, Zegura SL, and Hammer MF. 2002. High levels of Y-chromosome nucleotide diversity in the genus Pan. *Proc Natl Acad Sci U S A* **99:** 43-8.
Sundqvist AK, Ellegren H, Olivier M, and Vila C. 2001. Y chromosome haplotyping in Scandinavian wolves (Canis lupus) based on microsatellite markers. *Mol Ecol* **10:** 1959-66.

The International SNP Map Working Group. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409:** 928-33.
Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, and Goldstein DB. 1998. Origins of Old Testament priests. *Nature* **394:** 138-40.
Thomson R, Pritchard JK, Shen P, Oefner PJ, and Feldman MW. 2000. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc Natl Acad Sci U S A* **97:** 7360-5.
Tiepolo L, and Zuffardi O. 1976. Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. *Human Genetics* **34:** 119-124.
Toder R, Wakefield MJ, and Graves JA. 2000a. The minimal mammalian Y chromosome - the marsupial Y as a model system. *Cytogenet Cell Genet* **91:** 285-92.
Toder R, Wakefield MJ, and Graves JA. 2000b. The minimal mammalian Y chromosome - the marsupial Y as a model system. *Cytogenet Cell Genet* **91:** 285-92.
Toder R, Wienberg J, Voullaire L, O'Brien PC, Maccarone P, and Graves JAM. 1997. Shared DNA sequences between the X and Y chromosomes in the tammar wallaby - evidence for independent additions to eutherian and marsupial sex chromosomes. *Chromosoma* **106:** 94-8.
Tosi AJ, Morales JC, and Melnick DJ. 2003. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution Int J Org Evolution* **57:** 1419-35.
Troy CS, MacHugh DE, Bailey JF, Magee DA, Loftus RT, Cunningham P, Chamberlain AT, Sykes BC, and Bradley DG. 2001. Genetic evidence for Near-Eastern origins of European cattle. *Nature* **410:** 1088-91.
Trut LN. 1999. Early canid domestication: The farm fox experiment. *American Scientist* **87:** 160-169.
Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, and Oefner PJ. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* **7:** 996-1005.
Wallner B, Brem G, Muller M, and Achmann R. 2003. Fixed nucleotide differences on the Y chromosome indicate clear divergence between Equus przewalskii and Equus caballus. *Anim Genet* **34:** 453-6.
Van Hooft WF, Groen AF, and Prins HH. 2002. Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. *Mol Ecol* **11:** 267-79.
Wang RL, Stec A, Hey J, Lukens L, and Doebley J. 1999. The limits of selection during maize domestication. *Nature* **398:** 236-9.

Wardell BB, Sudweeks JD, Meeker ND, Estes SS, Woodward SR, and Teuscher C. 1993. The identification of Y chromosome-linked markers with random sequence oligonucleotide primers. *Mamm Genome* **4:** 109-12.

Waters PD, Duffy B, Frost CJ, Delbridge ML, and Graves JA. 2001. The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80-130 million years ago. *Cytogenet Cell Genet* **92:** 74-9.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420:** 520-62.

Watson JM, Spencer JA, Graves JA, Snead ML, and Lau EC. 1992. Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. *Genomics* **14:** 785-9.

Vila C, Leonard JA, Gotherstrom A, Marklund S, Sandberg K, Liden K, Wayne RK, and Ellegren H. 2001. Widespread origins of domestic horse lineages. *Science* **291:** 474-7.

Vila C, Walker C, Sundqvist AK, Flagstad O, Andersone Z, Casulli A, Kojola I, Valdmann H, Halverson J, and Ellegren H. 2003. Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity* **90:** 17-24.

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, and Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* **18:** 6531-5.

Vogel T, Borgmann S, Dechend F, Hecht W, and Schmidtke J. 1997. Conserved Y-chromosomal location of TSPY in Bovidae. *Chromosome Res* **5:** 182-5.

Vogel W, Jainta S, Rau W, Geerkens C, Baumstark A, Correa-Cerro LS, Ebenhoch C, and Just W. 1998. Sex determination in *Ellobius lutescens*: the story of an enigma. *Cytogenet Cell Genet* **80:** 214-221.

Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, Kohn FM, Schill WB, Farah S, Ramos C, Hartmann M, Hartschuh W, Meschede D, Behre HM, Castel A, Nieschlag E, Weidner W, Grone HJ, Jung A, Engel W, and Haidl G. 1996. Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* **5:** 933-43.

Wolf U, Flinspach G, Böhm R, and Ohno S. 1965. DNS-Reduplikationsmuster bei den riesen-geschlechtschromosomen von Microtus agrestis. *Chromosoma* **16:** 609-617.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, and et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23:** 4407-14.

Wyckoff GJ, Wang W, and Wu CI. 2000. Rapid evolution of male reproductive genes in the descent of man. *Nature* **403:** 304-9.

Acta Universitatis Upsaliensis

Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology*. (Prior to October, 1993, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science".)

Distribution:
Uppsala University Library
Box 510, SE-751 20 Uppsala, Sweden
www.uu.se, acta@ub.uu.se

ISSN 1104-232X
ISBN 91-554-5904-8