

Matching Multicity Y-STR Markers In Closely Related Individuals

Dipl.- Ing Thomas Krahn

Matching multicity Y-STR markers in closely related individuals - Multicity Y-STR markers are located on the symmetrical arms of palindromes. The DNA sequences on the palindromic arms are nearly identical and therefore highly prone to intrachromosomal recombination events. The talk will demonstrate how the recombination mechanisms will affect apparent mismatches in closely related individuals and it will present new tools that can help to understand the intrachromosomal re-arrangements in the Yq11 palindromic region.

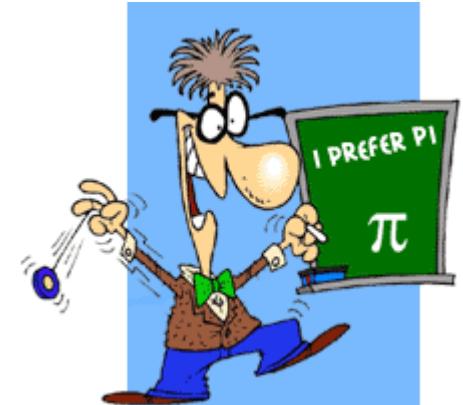
What's a palindrome?

Palindromic word phrases

RADAR

Madam, I'm Adam.
So many dynamos.
Cleveland DNA: Level C
Red rum, sir, is murder.
Too bad, I hid a boot.

Sat in a taxi, left Felix at Anita's
Emil asleep, Hannah peels a lime.
Anne, I stay a day at Sienna.
Was it a car or a cat I saw?



Max, I stay away at six a.m.
Go hang a salami! I'm a lasagna hog!
Yawn a more Roman way!

What's a palindrome?

Palindromic word phrases:

Forward = Backward

RADAR

Nucleic acids:

Forward = Reverse Complement (Forward)

AGCTTCTAGTCGACTAGAAAGCT

What's a palindrome?

Nucleic acids:

Forward = Reverse Complement (Forward)

AG**C**T**T**C**T**A**G**T**C**G**A**C**T**A**G**A**A**G**C**T

Reverse Complement (**A****G****C****T****T****C****T****A****G****T****C****G****A****C****T****A****G****A****G****C****T**)

= **A****G****C****T****T****C****T****A****G****T****C****G****A****C****T****A****G****A****G****C****T**

What's a palindrome?

Nucleic acids:

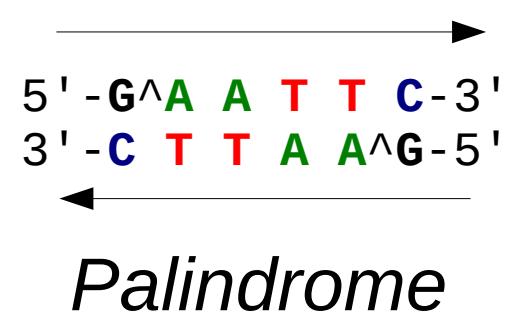
Forward = Reverse Complement (Forward)

AGCTTCTAGTCGACTAGAAAGCT

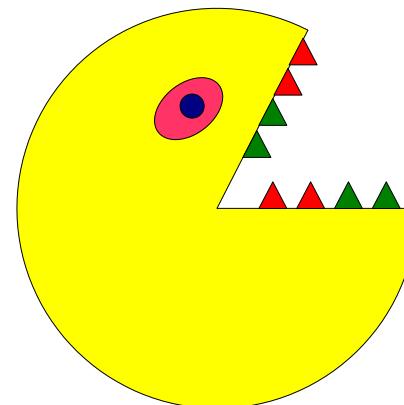


Biological Relevance of Palindromes

Restriction enzymes cut DNA at palindromic recognition sequences



EcoRI

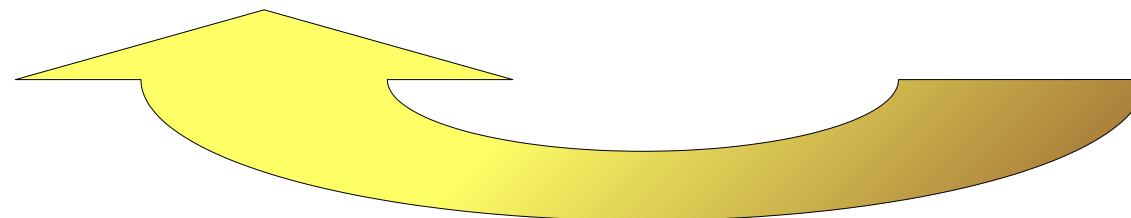


Biological Relevance of Palindromes

Formation of hairpins

Single stranded DNA with partial palindromic sequence

GAACTAGACTTAATGTGAGCTTATAAATTATAATGCTCTCACATTAAAGTCTAGTTC

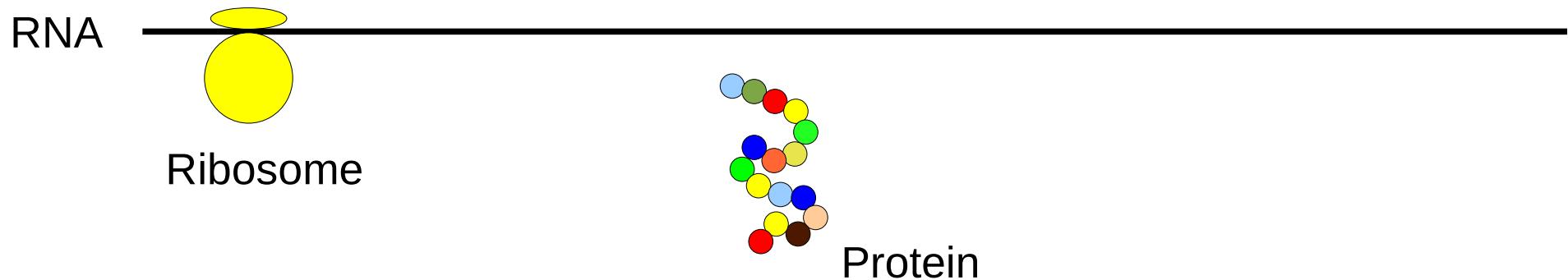


Hairpin structure

Biological Relevance of Palindromes

Hairpins regulate RNA translation

High temperature / low ionic concentration



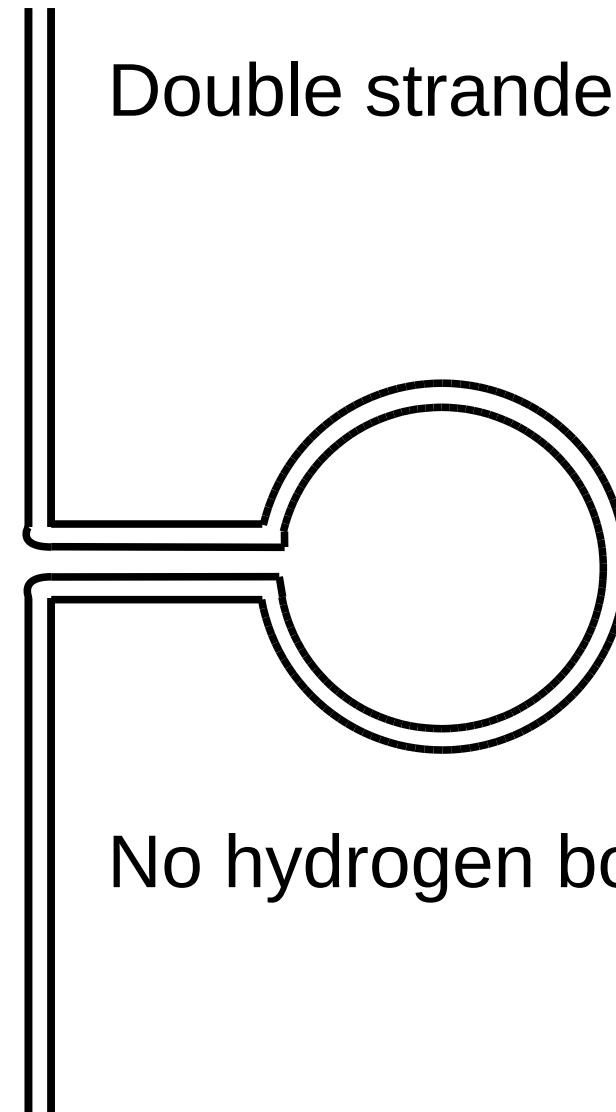
Low temperature / high ionic concentration



Palindromes in the ds Human Genome



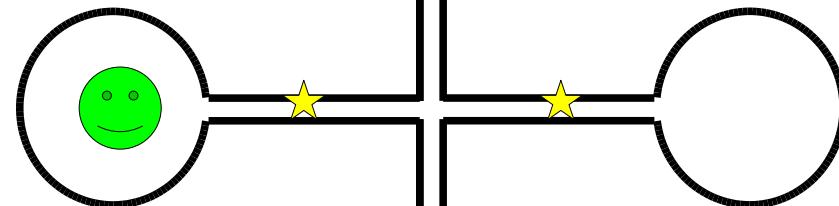
Chromosome



Palindromes in the ds Human Genome



Chromosome



But unexpected things happen...

Hydrogen bonding between
palindromic arms

Possible base differences will be
repaired by the cell's mutation
repair system

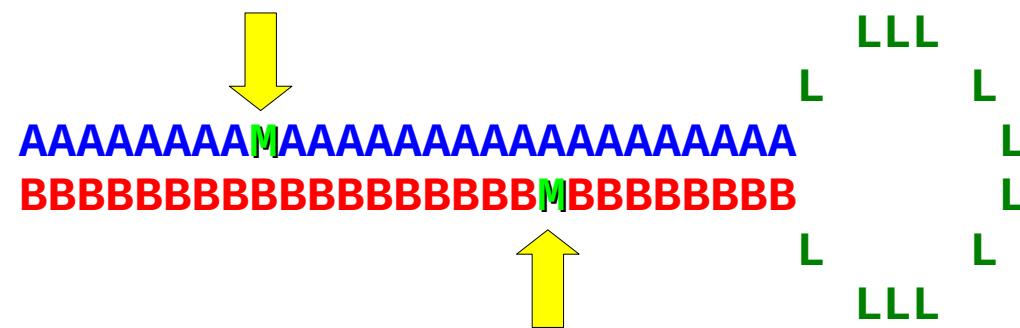
Recombinational Loss Of Heterozygosity (recLOH)

If we have a palindromic region (that means, that the DNA sequence has a loop and the ends of the loop can be put parallel next to each other and the parallel ends have nearly exactly the same DNA sequence, in other words an inverted repeat) we can model a simple structure like this:



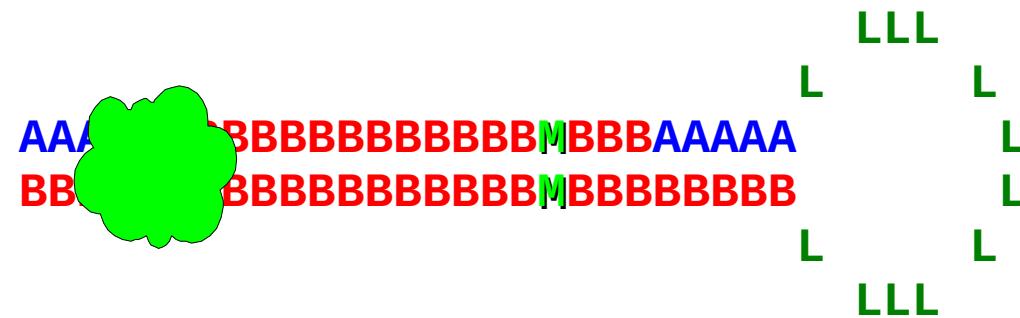
Recombinational Loss Of Heterozygosity (recLOH)

Palindromic sequences develop independently on each arm, so after some time they acquire mismatches (M) that could be STR mutations or SNPs on both arms. If it comes to a hairpin conformation, base pairing is not perfect and the cell tries to repair the apparent mutations.



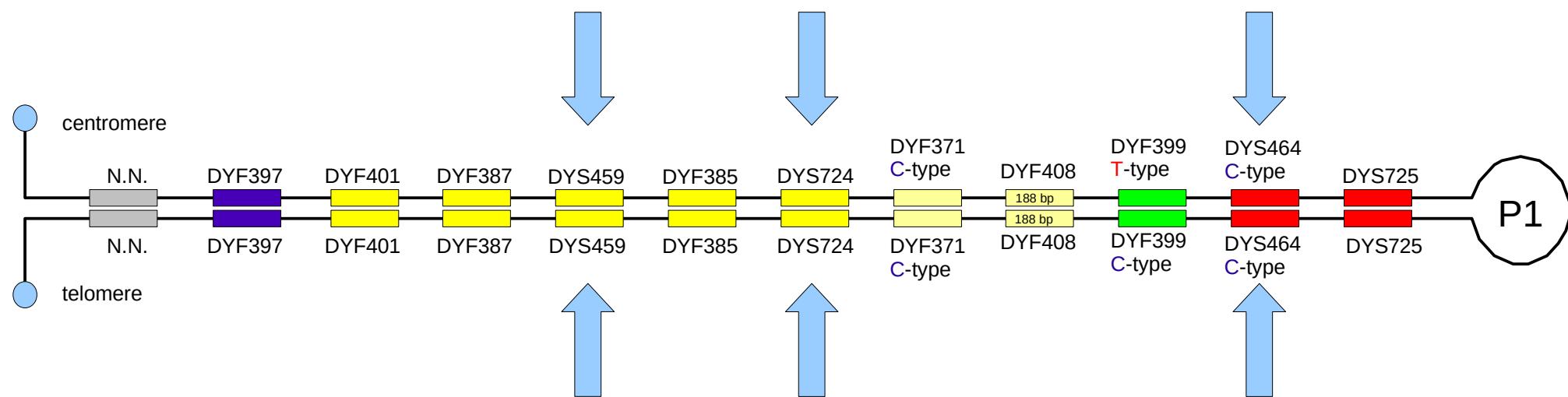
Recombinational Loss Of Heterozygosity (recLOH)

Depending on the direction of the repair enzyme complex, some mutations get duplicated on the other arm and some mutations disappear.



Recombinational Loss Of Heterozygosity (recLOH)

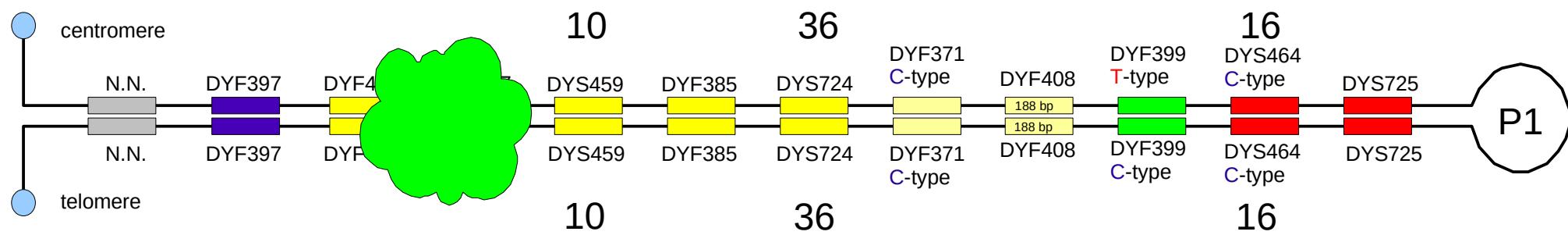
DYS459, DYS724 (CDY) and DYS464 are on the same palindrome called P1



Recombinational Loss Of Heterozygosity (recLOH)

In this example the haplotype has different alleles at all marker pairs

After the recombination event the heterozygosity is lost => recLOH



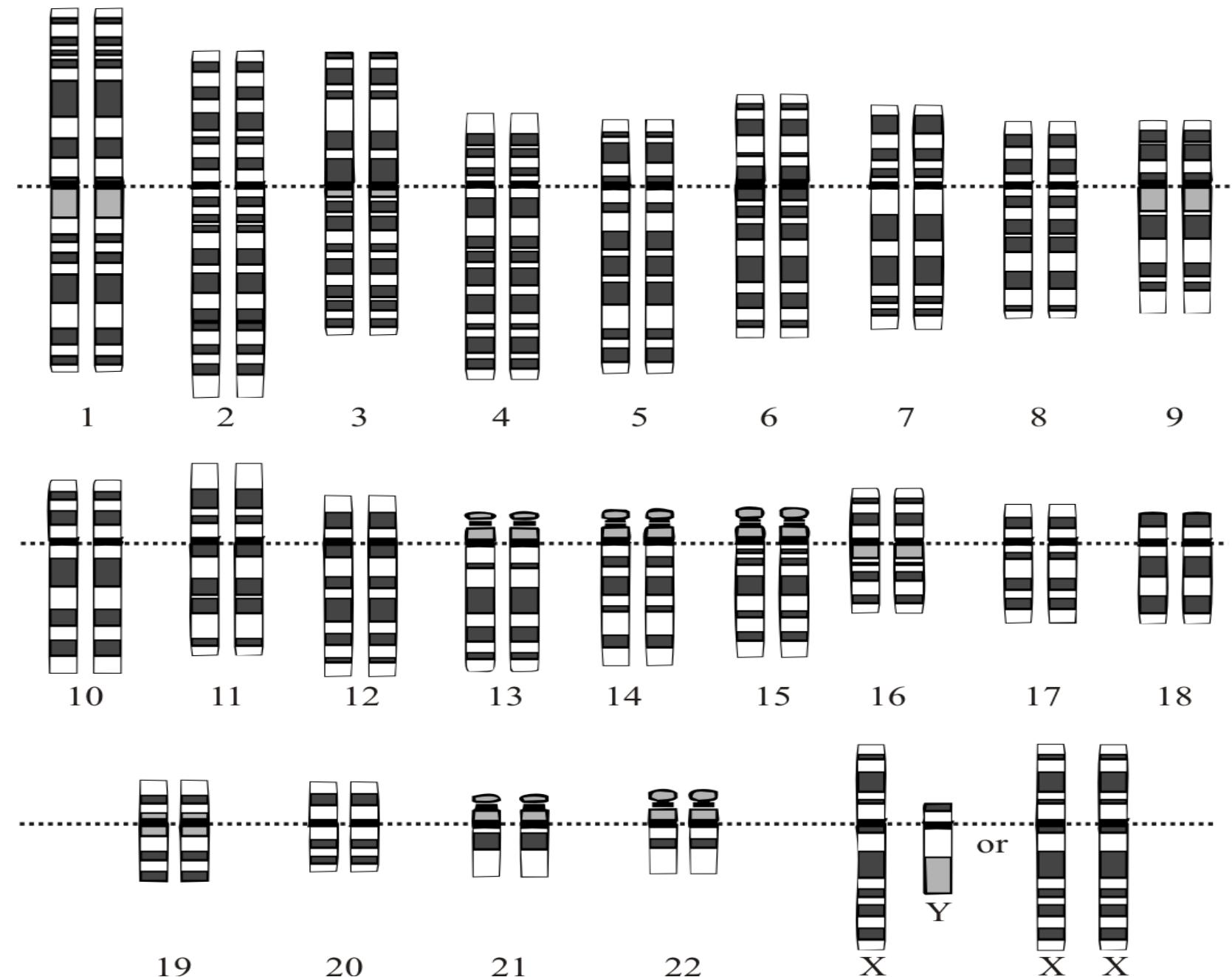
Typical recLOH Patterns In Genetic Genealogy

A classical DYS459, CDY, DYS464 recLOH

DYS459	DYS724 (CDY)	DYS464	Genetic distance?
9-10	37-38	14-14-16-18	
9-10	37-38	14-15-16-18	1
9-10	37-38	14-14-16-18	0
9-10	37-39	14-14-16-18	1
9-10	37-38	14-14-16-18	0
10-10	38-38	14-14-18-18	1
9-9	37-37	14-16-18-18	1

Similar recLOH events happen at all palindromic multicopy markers

DYS385	YCAII
11-14	19-23
11-14	19-23
11-14	19-19
11-11	19-23



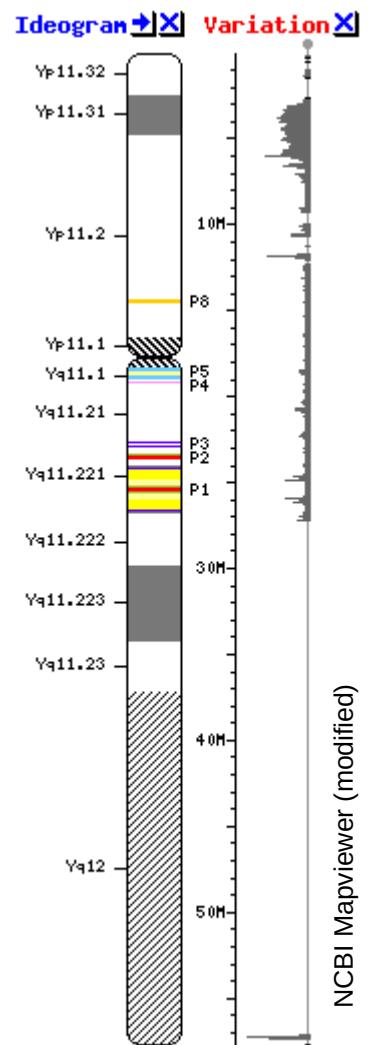
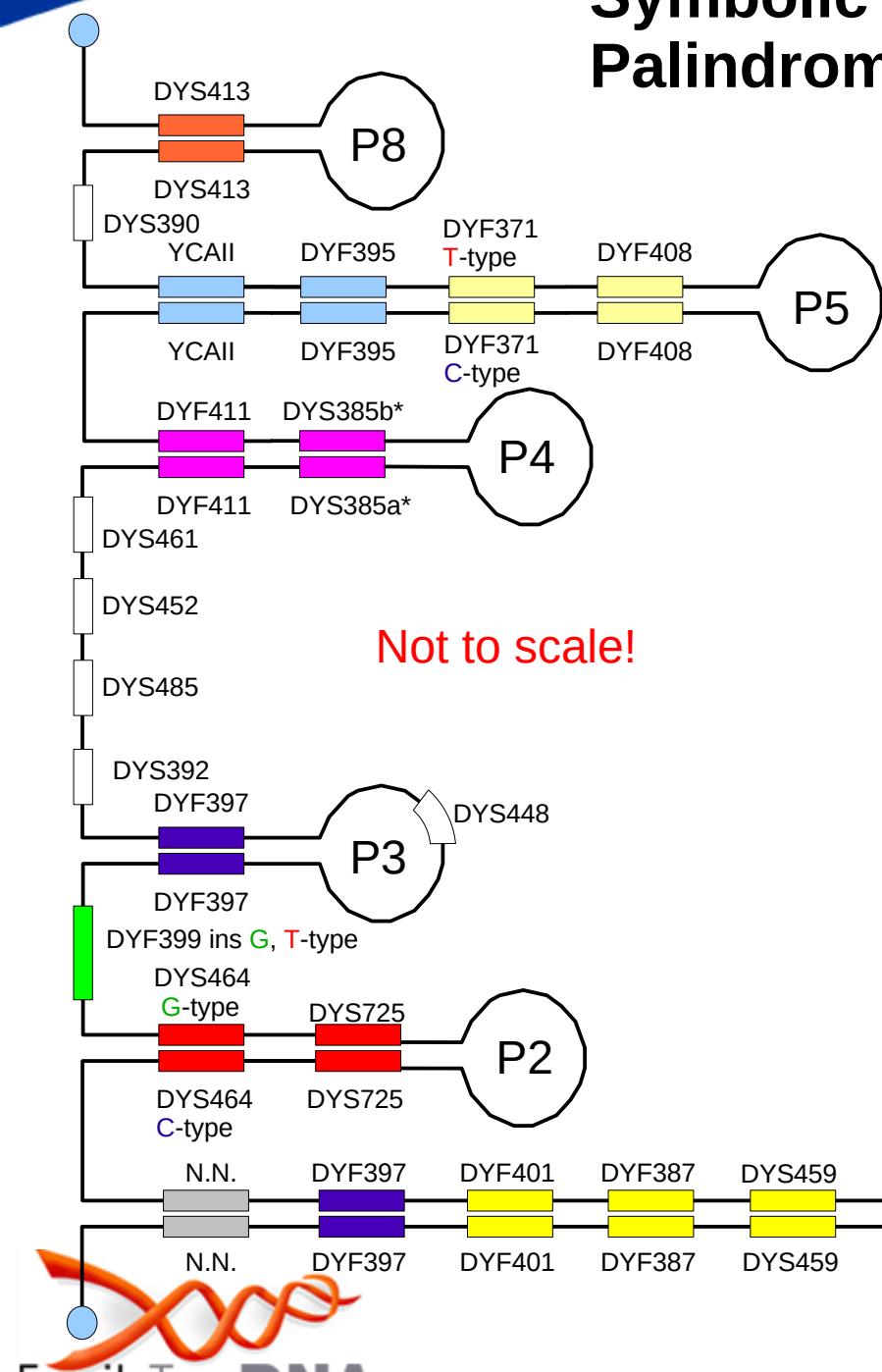
Highest Density of Palindromes in the Human Genome

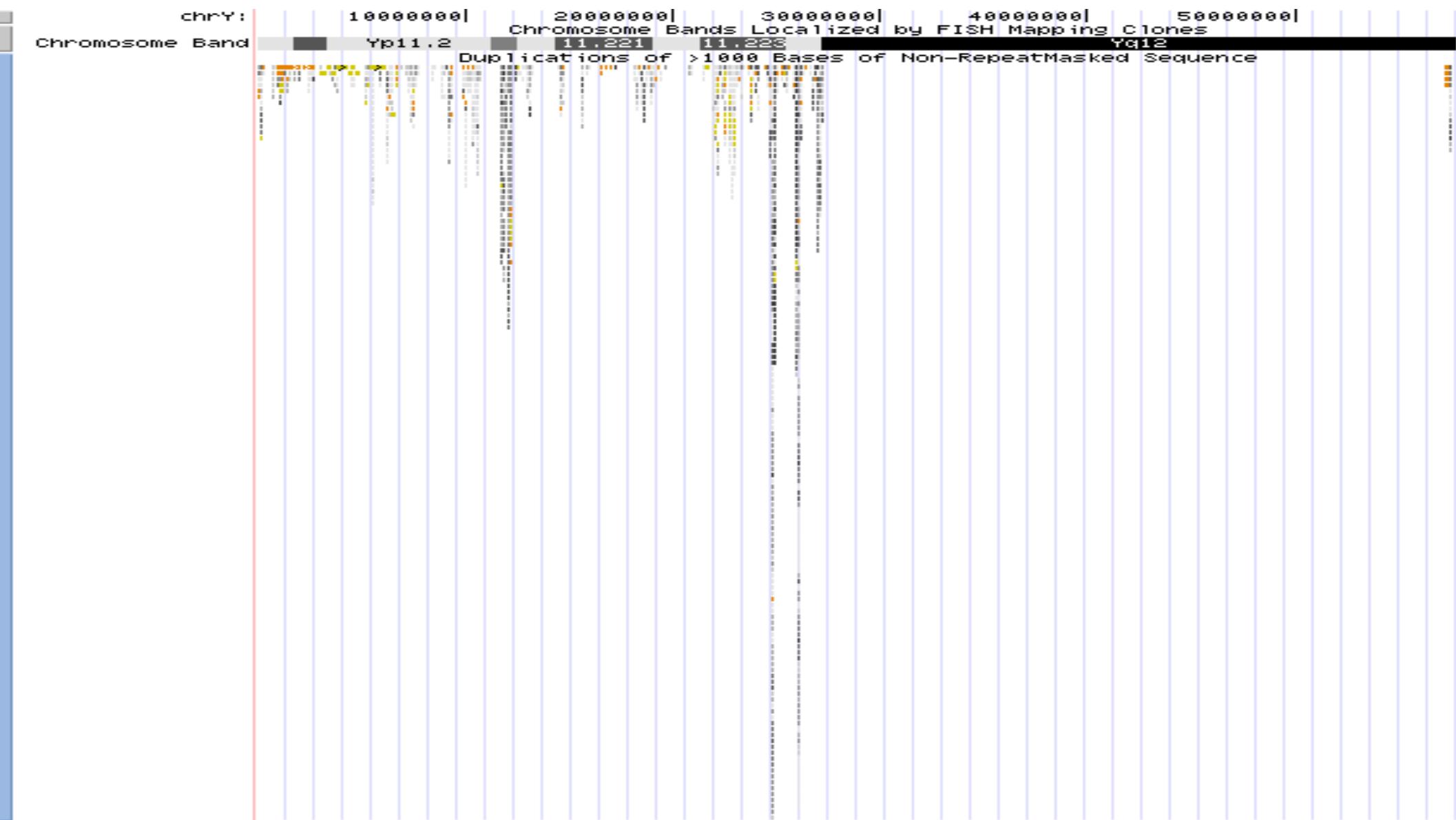
Segmental Duplications in the Human Genome

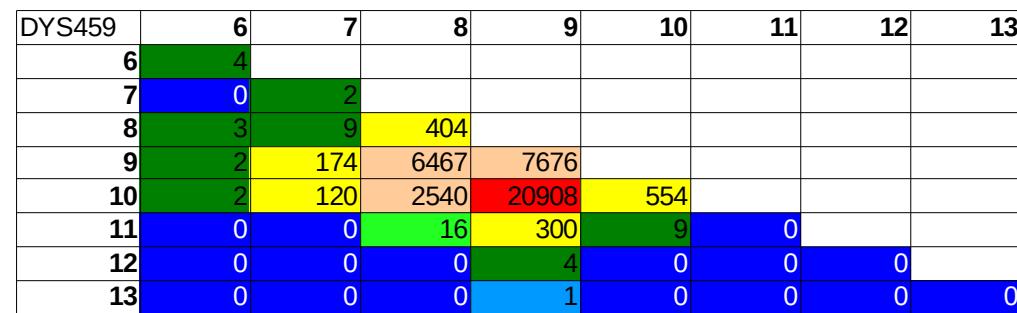
Chr	Length (bp)	Intrachromosomal		Interchromosomal		Total	
		Length	%	Length	%	Length	%
1	245,203,898	6,431,462	2.6	3,964,057	1.6	8,678,912	3.5
2	243,315,028	6,380,301	2.6	3,757,634	1.5	8,935,221	3.7
3	199,411,731	1,646,046	0.8	1,870,056	0.9	2,671,459	1.3
4	191,610,523	2,323,764	1.2	2,547,466	1.3	3,927,792	2.0
5	180,967,295	4,066,897	2.2	2,083,920	1.2	5,208,550	2.9
6	170,740,541	2,048,892	1.2	1,123,050	0.7	2,854,222	1.7
7	158,431,299	9,629,716	6.1	3,734,503	2.4	11,722,991	7.4
8	145,908,738	1,576,863	1.1	1,694,593	1.2	2,153,612	1.5
9	134,505,819	8,451,476	6.3	4,371,262	3.2	9,403,888	7.0
10	135,480,874	6,460,047	4.8	1,919,342	1.4	7,741,228	5.7
11	134,978,784	4,223,832	3.1	2,147,666	1.6	5,382,256	4.0
12	133,464,434	1,616,743	1.2	1,134,900	0.9	2,582,114	1.9
13	114,151,656	1,451,225	1.3	1,655,399	1.5	2,700,321	2.4
14	105,311,216	282,478	0.3	849,400	0.8	1,116,676	1.1
15	100,114,055	5,520,203	5.5	3,339,498	3.3	7,091,918	7.1
16	89,995,999	7,378,691	8.2	3,456,338	3.8	8,247,312	9.2
17	81,691,216	5,505,106	6.7	1,217,149	1.5	6,432,722	7.9
18	77,753,510	230,844	0.3	1,400,896	1.8	1,627,497	2.1
19	63,790,860	1,763,189	2.8	918,571	1.4	2,531,577	4.0
20	63,644,868	772,190	1.2	1,068,246	1.7	1,369,456	2.2
21	46,976,537	431,633	0.9	1,714,574	3.6	1,734,567	3.7
22	49,476,972	2,303,175	4.7	1,633,388	3.3	3,481,523	7.0
Y	152,634,166	3,579,325	2.3	4,550,908	3.0	8,047,172	5.3
	50,961,097	6,651,452	13.1	1,462,582	2.9	7,353,078	14.4
Total	3,070,521,116	90,725,550	3.0	53,615,398	1.7	122,996,064	4.0

Liqing Zhang et al. "Patterns of Segmental Duplication in the Human Genome", Mol. Biol. Evol. 22(1):135–141. 2005

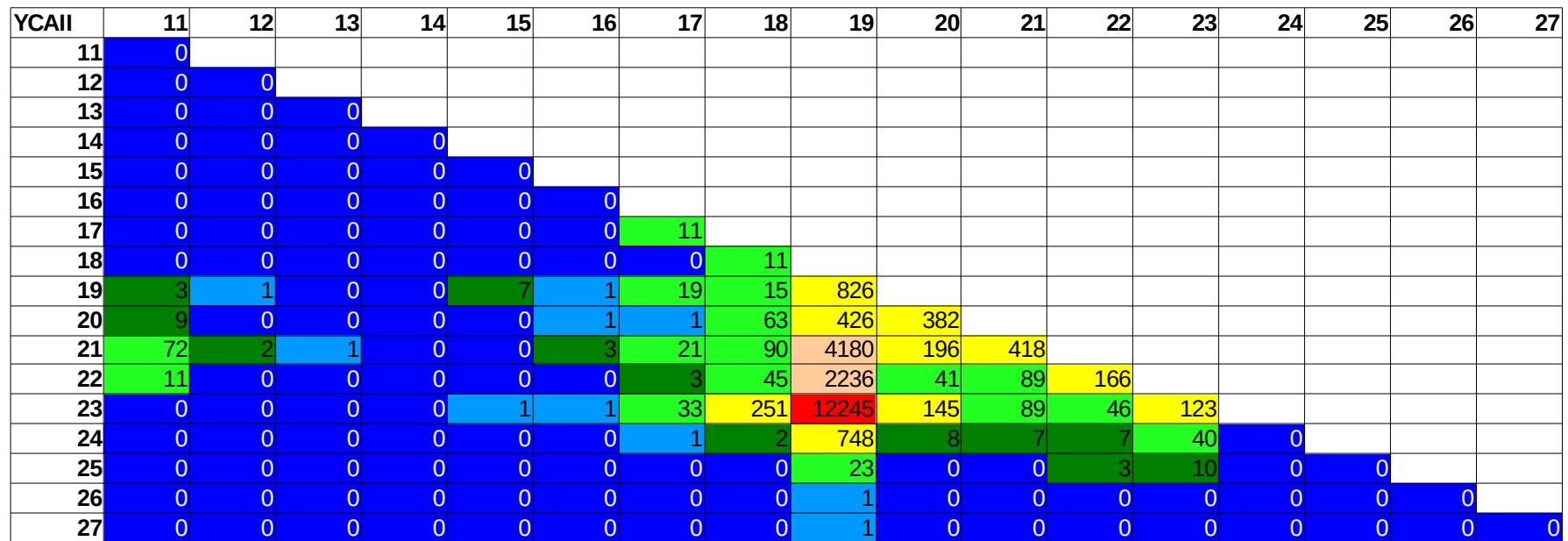
Symbolic Map of the Yq11 Palindromic Region (Version 2)



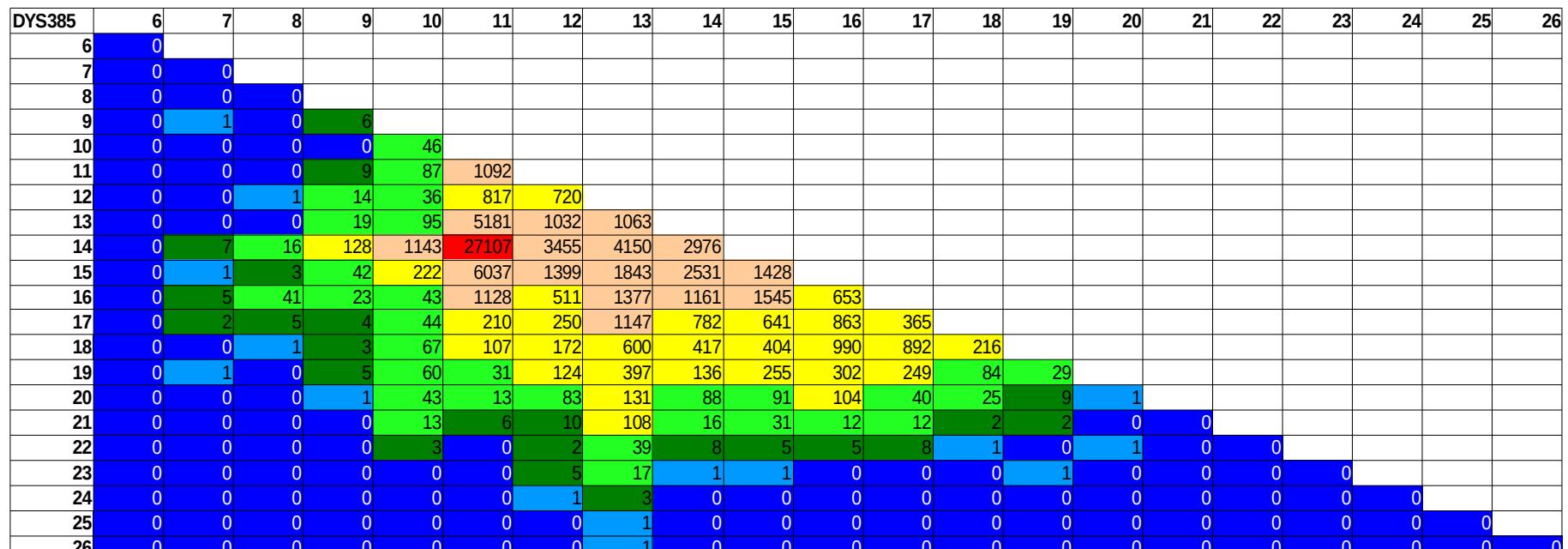




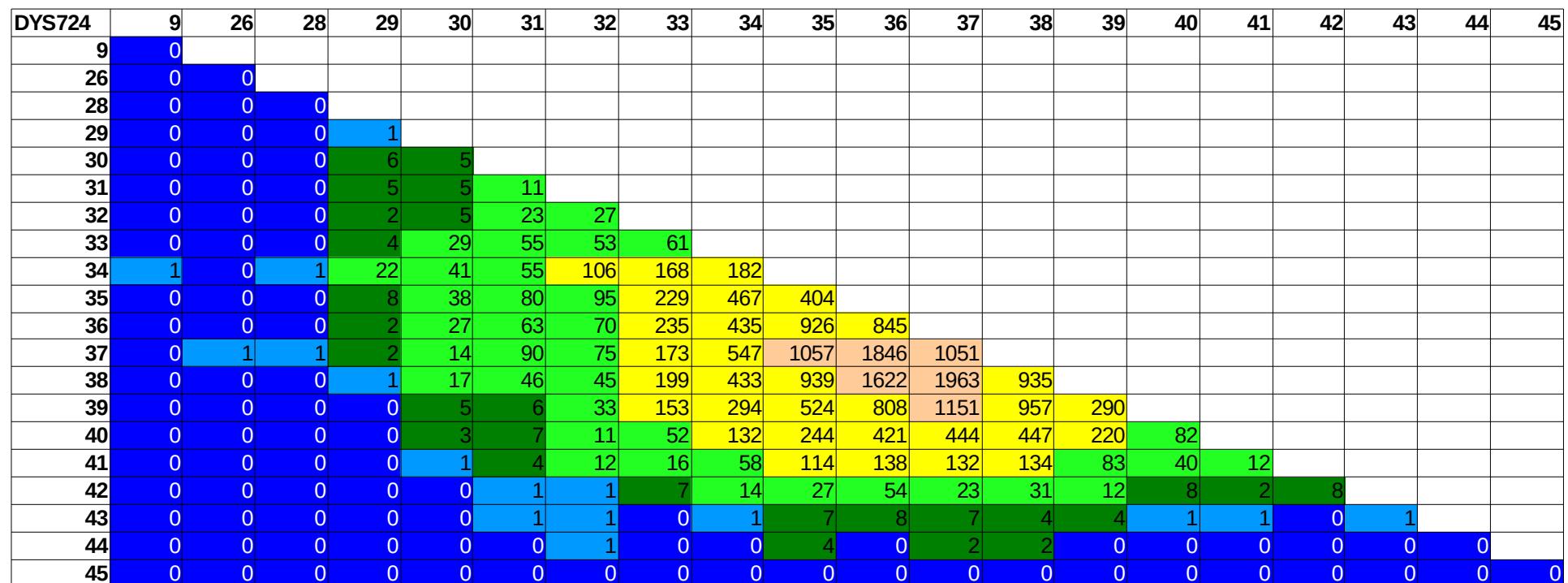
Scale: 0 1 10 100 1000 10000 100000



Scale: 0 1 10 100 1000 10000 100000

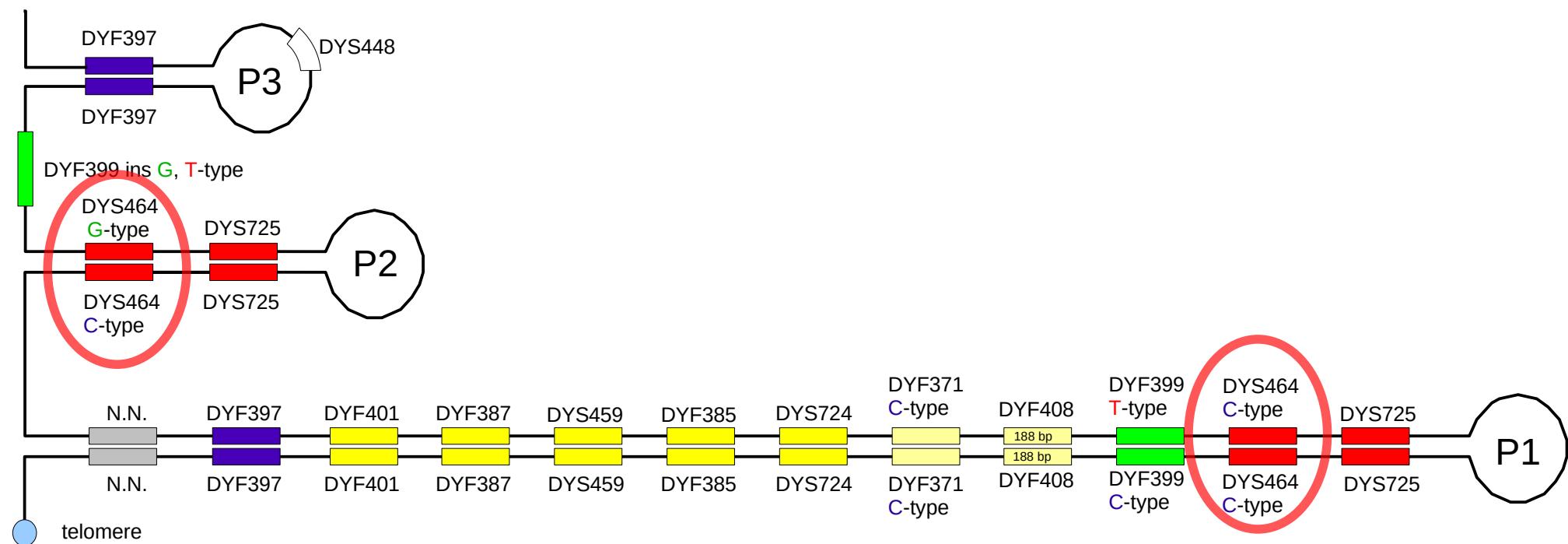


Scale: 0 1 10 100 1000 10000 100000

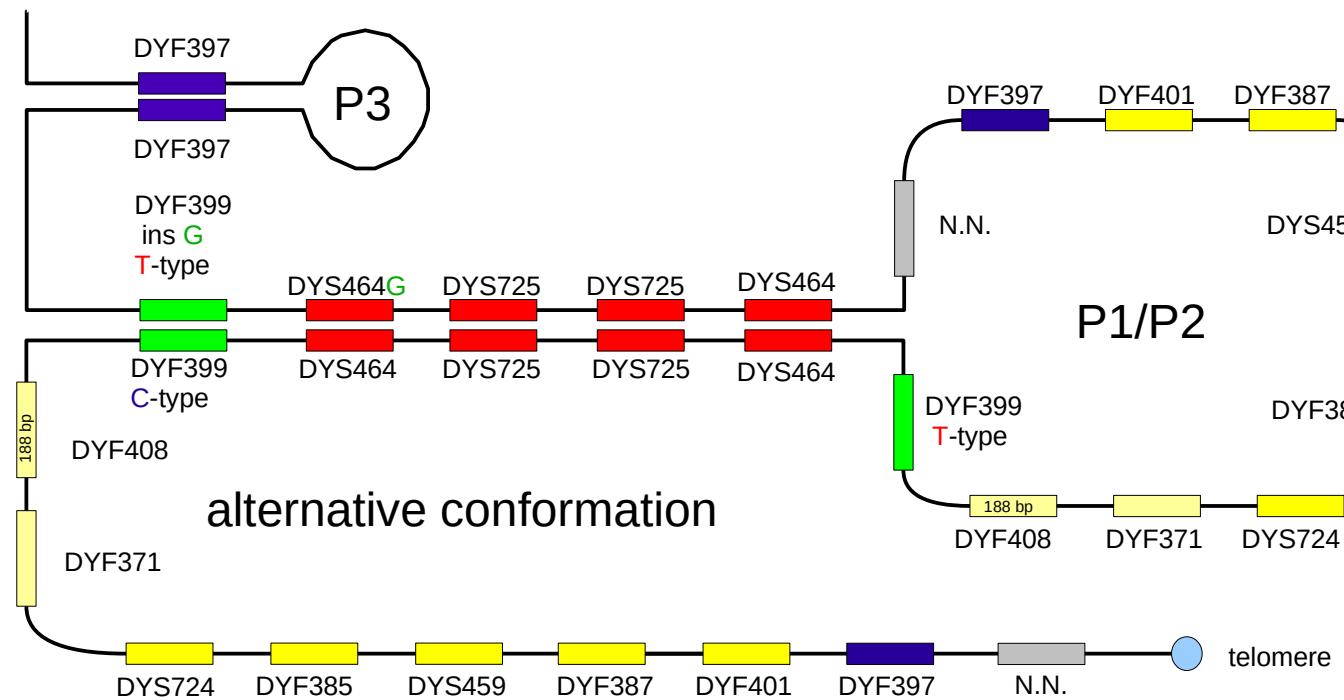


Scale: 0 1 10 100 1000 10000

DYS464 – A Complex Marker With (Usually) 4 Alleles



DYS464 – A Complex Marker With (Usually) 4 Alleles



DYS464 Extended PCR

Fluorescein

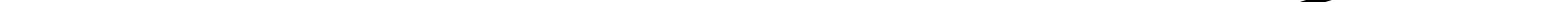
AAGAGTCTCTGTACCCAGG
t gaaaagagtctctgttacccagg
t gaaaagagtctctgttacccagg

ccttcctc~~t~~ttcttctttttttccctgaaagagtgctctgttacccaggatggtgtat

1

1

1

JOE

8



JOE



TAMRA



DYS464 Extended PCR

If the primer ending with **G** yields a **peak**, the allele is classed as a **G-type**.

If the primer ending with **C** yields a **peak**, the allele is classed as a **C-type**.

DYS464-forward

FAM-ttacgagcttggctatg (exactly like Redd paper)

DYS464-reverse-G-Type

JOE-cctggtaacagagagactcttcag (ending with G)

DYS464-reverse-C-Type

TMR-cctggtaacagagagactcttcac (ending with C)

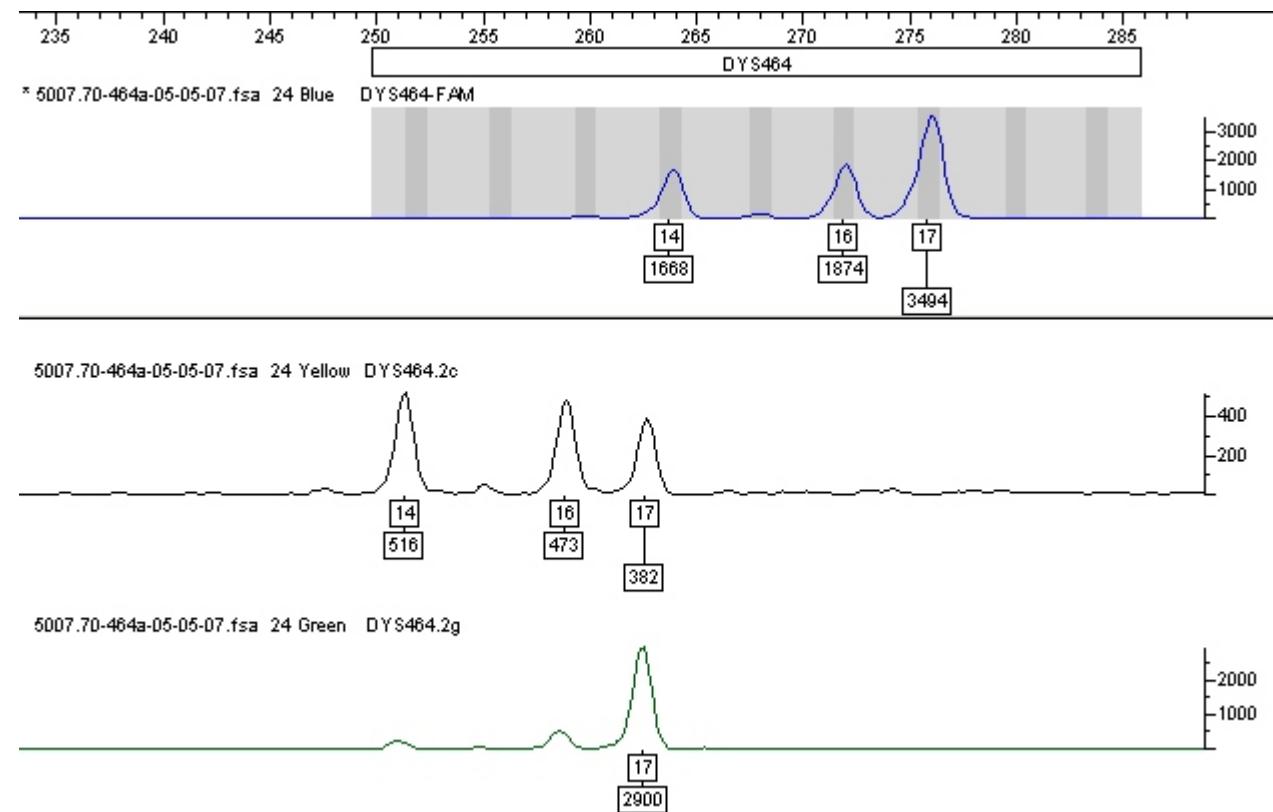
PCR parameters: 30 x (90°, 65°, 70°) and 60° end-elongation 1/2 h

Peaks with a **green fluorescence** (**JOE**) are called **G-type**.

Peaks with a **yellow fluorescence** (**TMR**) are called **C-type**.

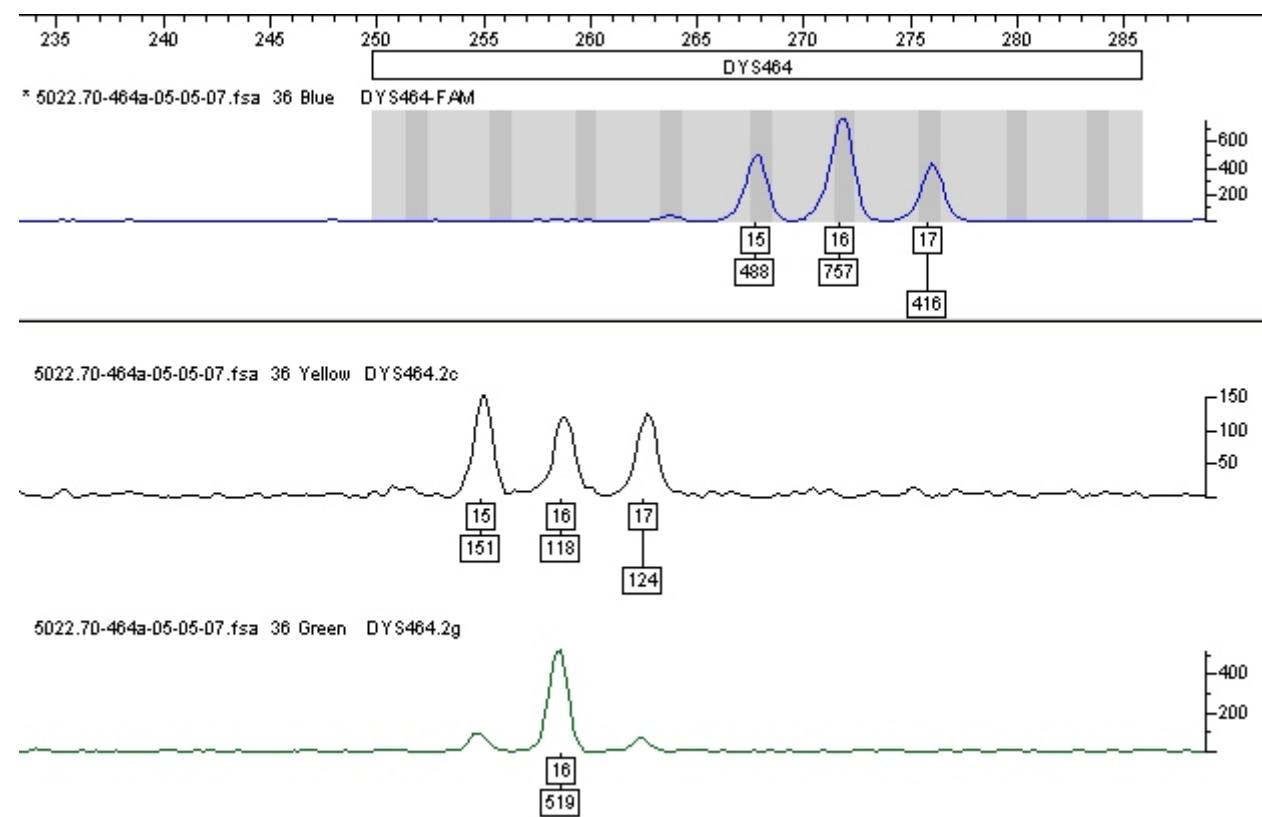
DYS464 Extended PCR

Electropherogram



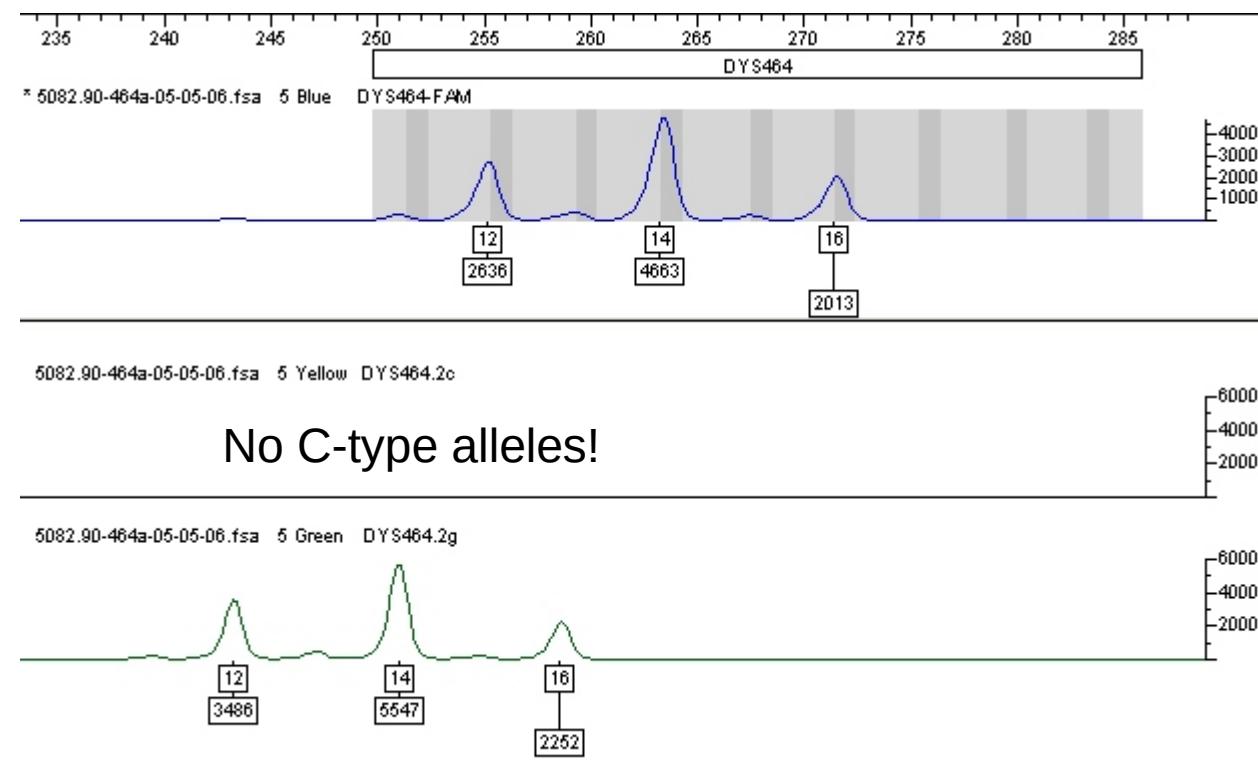
DYS464 Extended PCR

Electropherogram



DYS464 Extended PCR

Haplogroup I



Typing of DYS464X

Y haplogroup	DYS 464
A	1 1 g -1 3 g -1 3 g -1 6 g
E	1 4 g -1 5 .3 g -1 7 g -1 8 g
E3b1	1 4 g -1 5 .3 g -1 7 g -1 8 g
G	1 3 g -1 4 g -1 5 g -1 5 g
G2*	1 2 g -1 2 g -1 2 g -1 3 g
I	1 2 g -1 4 g -1 5 g -1 6 g
I1a	1 2 g -1 4 g -1 4 g -1 6 g
I1a3	1 2 g -1 2 g -1 4 g -1 4 g -1 5 g -1 6 g
I1b	1 1 g -1 4 g -1 4 g -1 4 g
I1b	1 1 g -1 4 g -1 4 g -1 5 g
I1b2a	1 1 g -1 4 g -1 4 g -1 5 g
I1b2a1	1 1 g -1 1 g -1 4 g -1 5 g
I1c	1 4 g -1 5 g -1 5 g -1 6 g
J2a1*	1 2 g -1 3 g -1 5 g -1 6 g -1 6 g -1 6 g
N	1 4 g -1 4 .3 g
R1a1*	1 2 g -1 5 g -1 5 g -1 6 g
R1b	1 6 c -1 6 c -1 6 g -1 6 g
R1b	1 5 c -1 6 c
R1b	1 5 c -1 5 c -1 7 c -1 7 g
R1b	1 4 c -1 6 c -1 7 c -1 7 g
R1b	1 4 c -1 5 c -1 6 g -1 7 c
R1b	1 6 c -1 6 g
R1b	1 5 c -1 5 c -1 6 c -1 6 g
R1b	1 4 c -1 5 c -1 7 c -1 7 g
R1b	1 5 c -1 6 c -1 7 g -1 7 g
R1b	1 5 c -1 6 c
R1b	1 5 c -1 5 c -1 5 c -1 5 c
R1b	1 5 c -1 5 c -1 7 c -1 7 g
R1b	1 5 c -1 7 c -1 7 c -1 8 g
R1b	1 5 c -1 5 c -1 7 c -1 8 g
R1b	1 5 c -1 5 c -1 6 g -1 7 c
R1b	1 6 c -1 6 c -1 7 c -1 7 g
R1b	1 4 c -1 5 c -1 5 c -1 5 g
R1b	1 5 c -1 5 c -1 6 c -1 8 g
R1b	1 5 c -1 5 c -1 6 c -1 7 .1 g
R1b	1 5 c -1 5 c -1 6 g -1 7 c
R1b	1 3 c -1 5 c -1 7 c -1 7 g
R1b	1 5 c -1 5 c -1 7 g -1 7 g
R1b	1 5 c -1 6 c -1 6 c -1 8 g
R1b	1 5 c -1 5 c -1 6 c -1 7 c

Other haplogroups have
only G-type alleles

R1b has usually 3 C-type
alleles and one G-type allele

Exceptions most likely
products of recLOH

What Is The Use Of The DYS464X Test?

Most R1b males that we have tested show 3 C-types and 1 G-type.
All other haplogroups (including R1a) show 4 G-types.

- 1.) The DYS464X method can verify, if a person is really R1b
- 2.) In many cases we can determine, if a single large peak really represents two different alleles
- 3.) Some DYS464 patterns look similar when the conventional test is used, but really consist of completely different alleles. For example:

14**c**-15**c**-17**c**-17**g** is completely different from
14**c**-15**g**-17**c**-17**c**,

but this wouldn't be seen with the conventional typing method.

- 4.) DYS464X is a method that helps us observe recombinational loss of heterozygosity (recLOH)

DYS464 Examples From An Arbitrary Surname Project:

(conventional results)

surname A, 1: 14-14-16-17

surname A, 2: 14-14-16-17

surname A, 3: 14-14-16-17

surname B, 1: 14-14-16-18

surname C, 1: 14-14-15-17

If all the 5 persons are R1b then they can be typed with an extended DYS464 test.

We might get the following results:

surname A, 1: 14c-14c-16g-17c

surname A, 2: 14c-14c-16g-17c

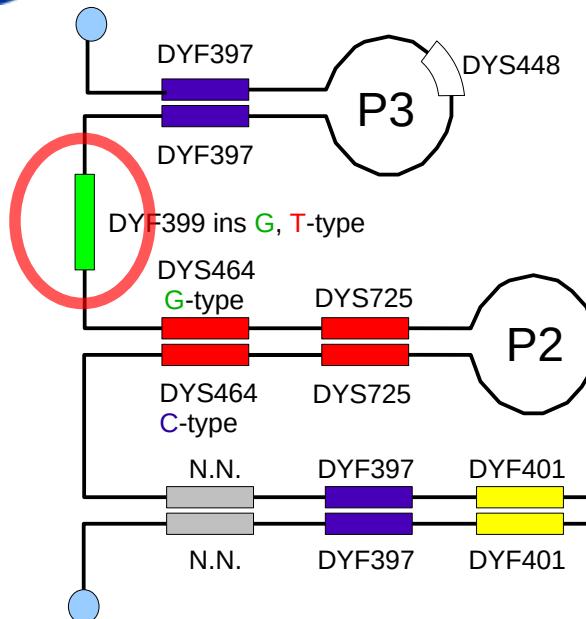
surname A, 3: 14c-14c-16c-17g

surname B, 1: 14c-14c-16g-18c

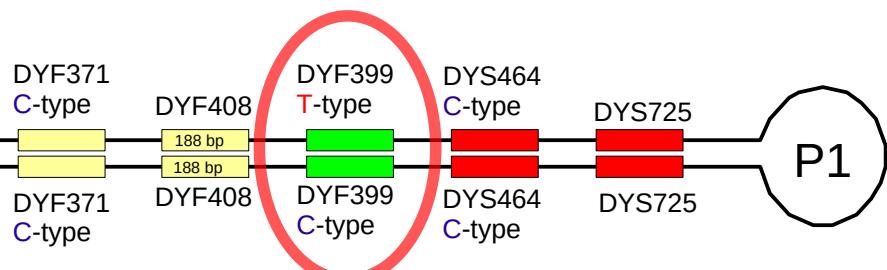
surname C, 1: 14c-14c-15g-17c

From that result we could say:

- A1 and A2 match completely
- A3 has at least a genetic distance of 2 from (A1 & A2)
- B1 has a genetic distance of 1 from (A1 & A2)
- C1 has a genetic distance of 1 from (A1 & A2)



DYF399 A Fast Moving, Asymmetrical Palindromic Y-STR



ins G

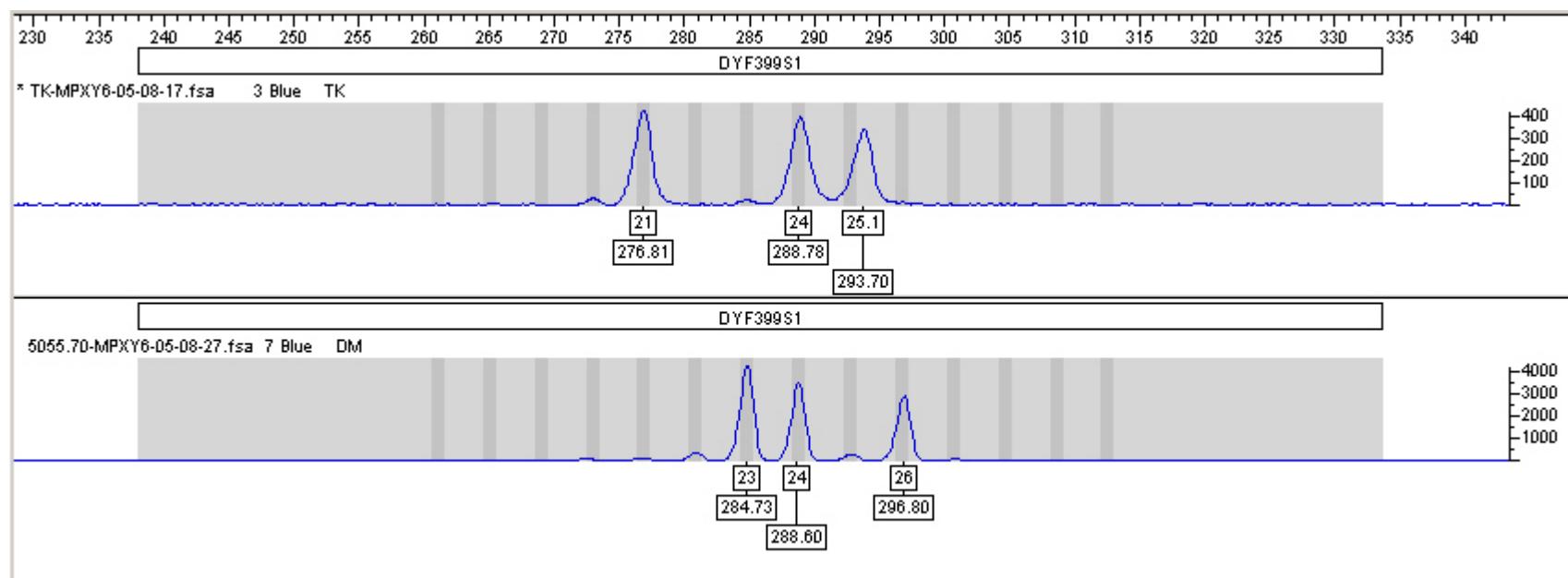
poly A

C-type and T-type reverse primers

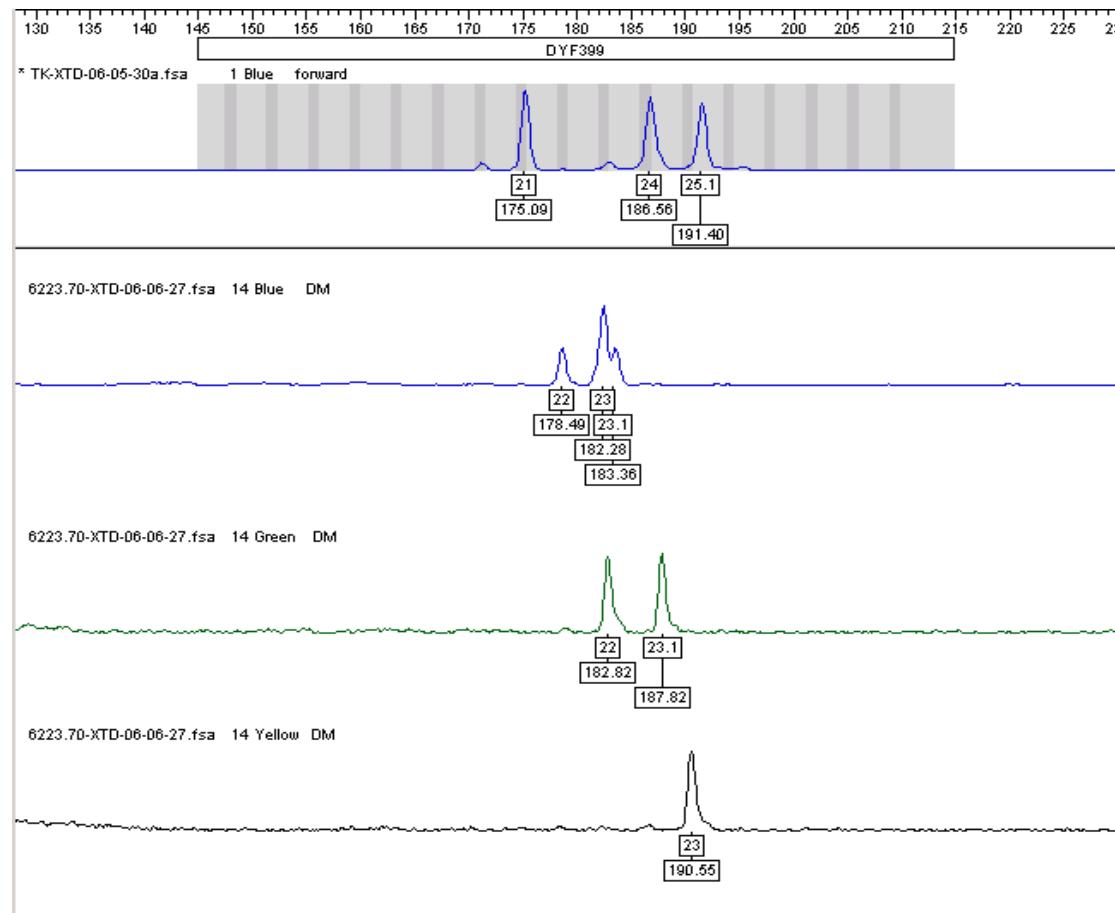


DYF399

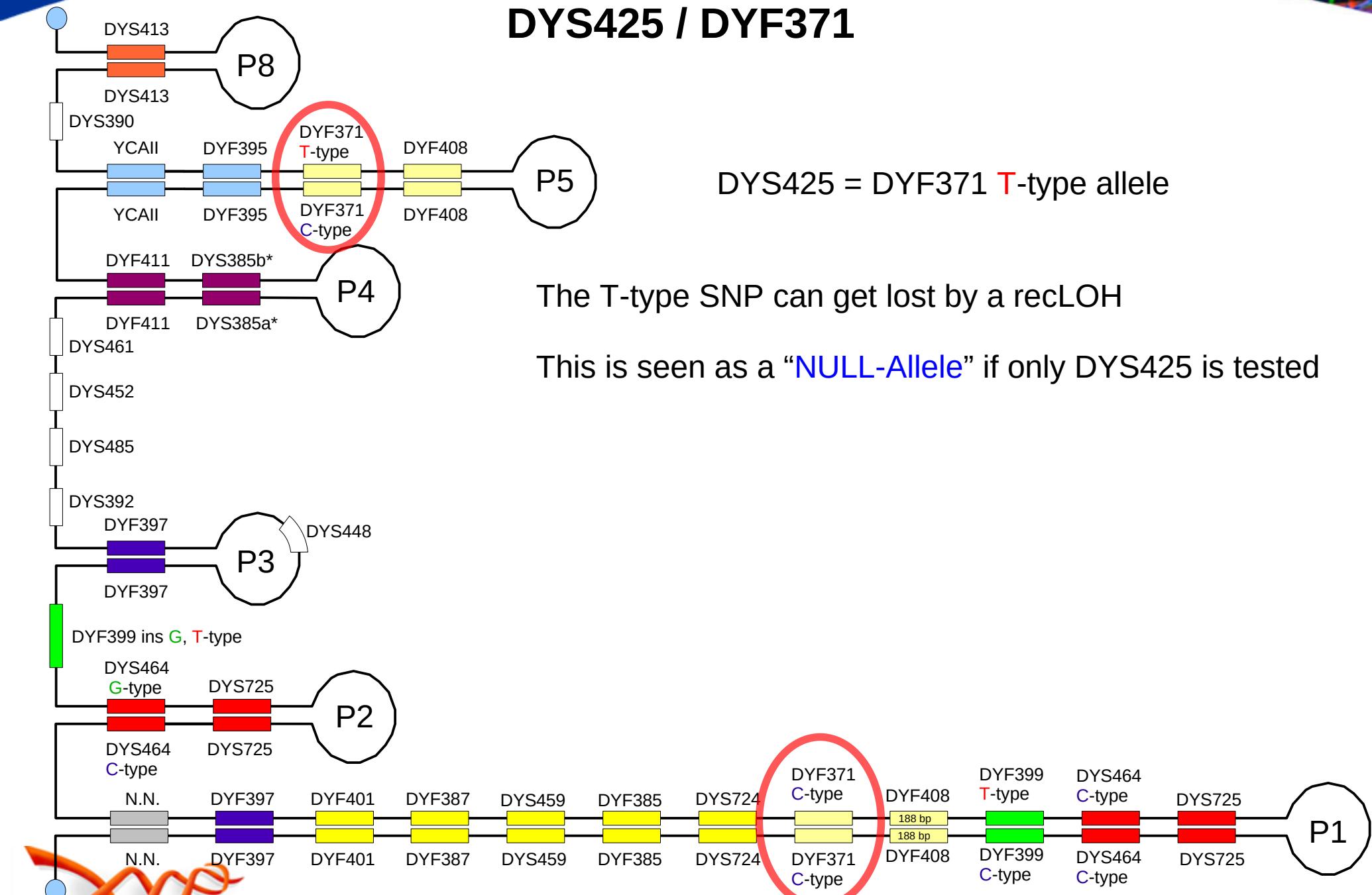
A Fast Moving, Asymmetrical Palindromic Y-STR



DYF399 - A Fast Moving, Asymmetrical Palindromic Y-STR



DYS425 / DYF371



DYS425 / DYF371

The HUGO sequence has also a Null allele at DYS425

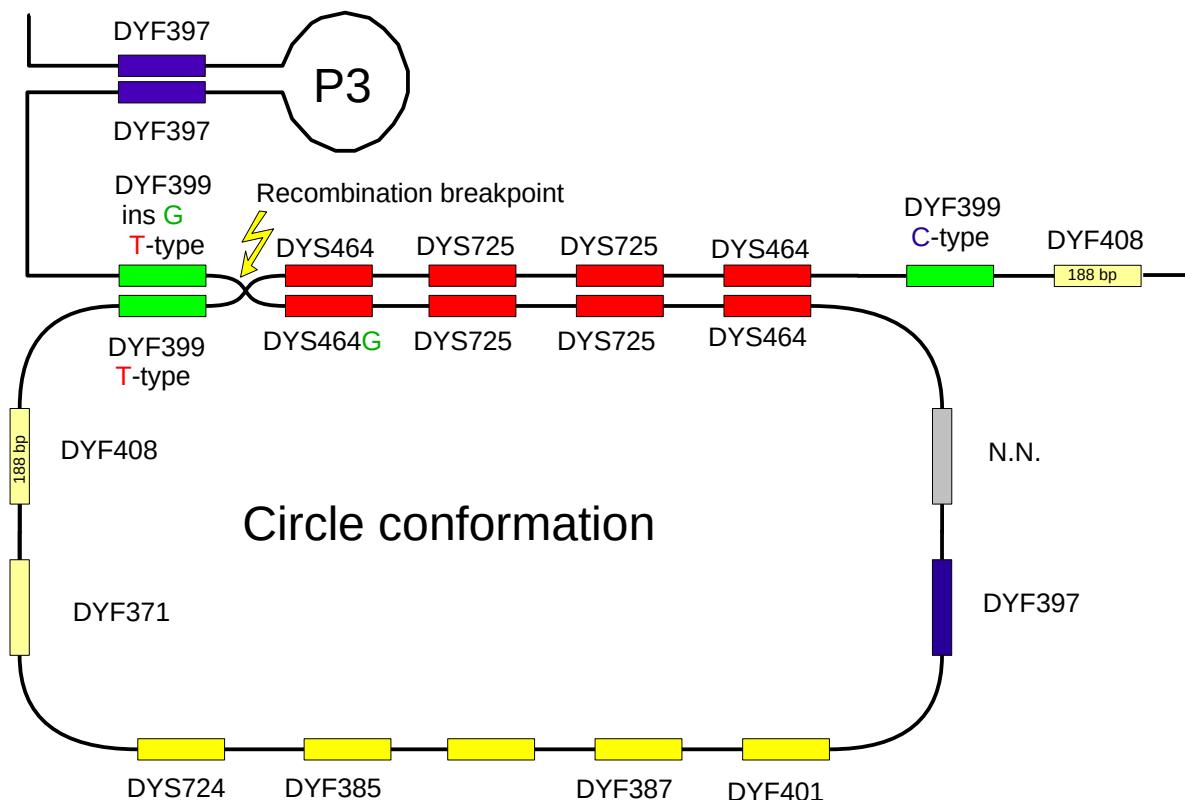
10c-10c-13c-14c

Normally in R1b (and most other haplogroups):

10c-12t-13c-14c

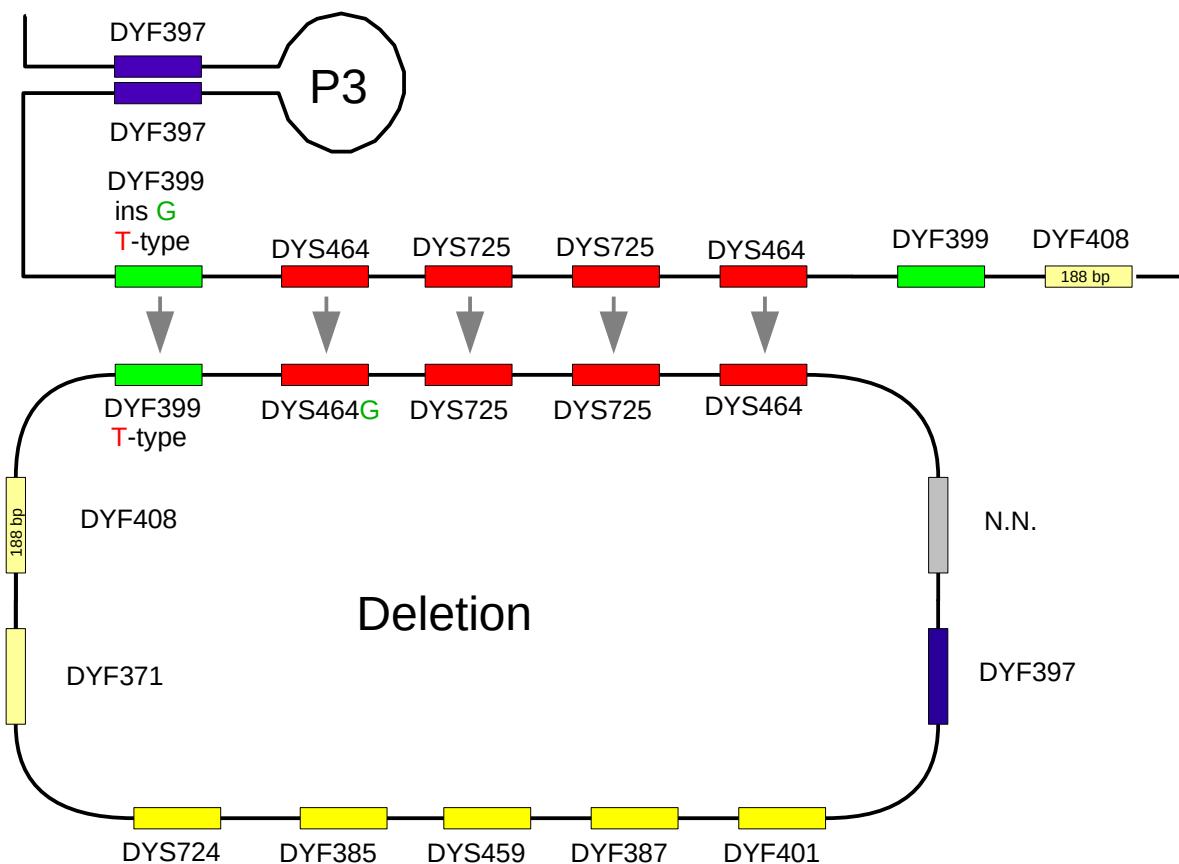
How comes It To A Deletion?

Symetry in the red/red (P1P2) region allows another irregular conformation:

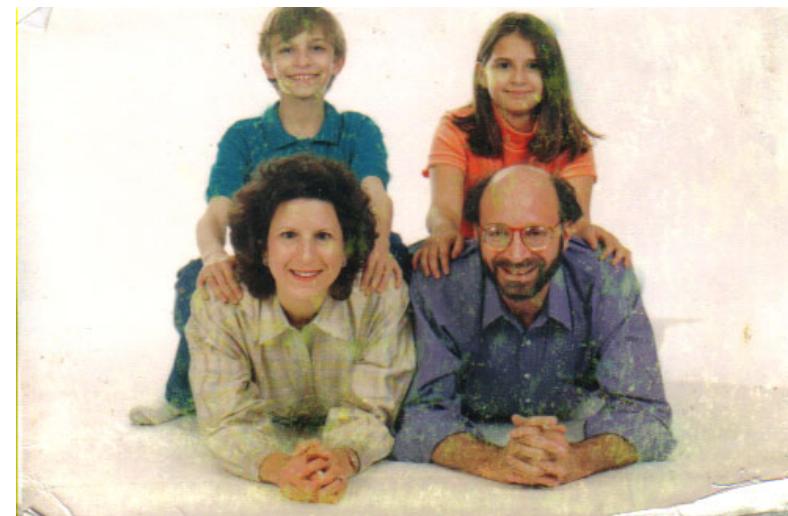


How comes It To A Deletion?

The circular DNA molecule can't replicate on its own and gets lost in the next cell cycle



Famous People with a P1/P2 Deletion



Symbolic Map of the Yq11 Palindromic Region (Version 2)

