

grass genomics: how to portion a mammoth?



david kopecký

laboratory of molecular cytogenetics and cytometry
centre of the region haná for biotechnological
and agricultural research
olomouc, czech republic

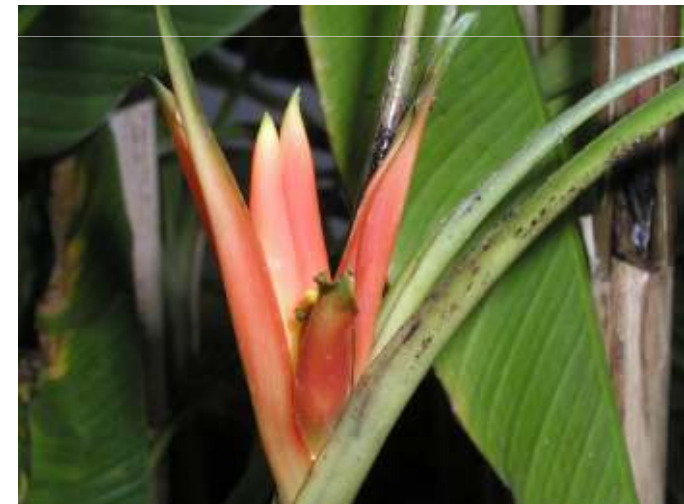
Olomouc Research Centre

- Two laboratories:
 - Laboratory of Molecular Cytogenetics and Cytometry
(Dr. J. Doležel)
 - Laboratory of Plant Cytoskeleton and Cell Cycle
(Dr. P. Binarová)
- 38 employees
 - 17 PhD.
- Collaboration with Faculty of Science and Pedagogical Faculty
(Palacky University in Olomouc)
 - BSc., MSc. a PhD. Students
 - lectures and practices



Research Topics

- Structure and evolution of plant genomes
- Our experimental subjects:
 - Bananas and Plantains (*Musa* spp.)
 - Festulolium (*Festuca* x *Lolium* hybrids)
 - Cereals (wheat, barley, rye, maize)



Grasses with improved characteristics

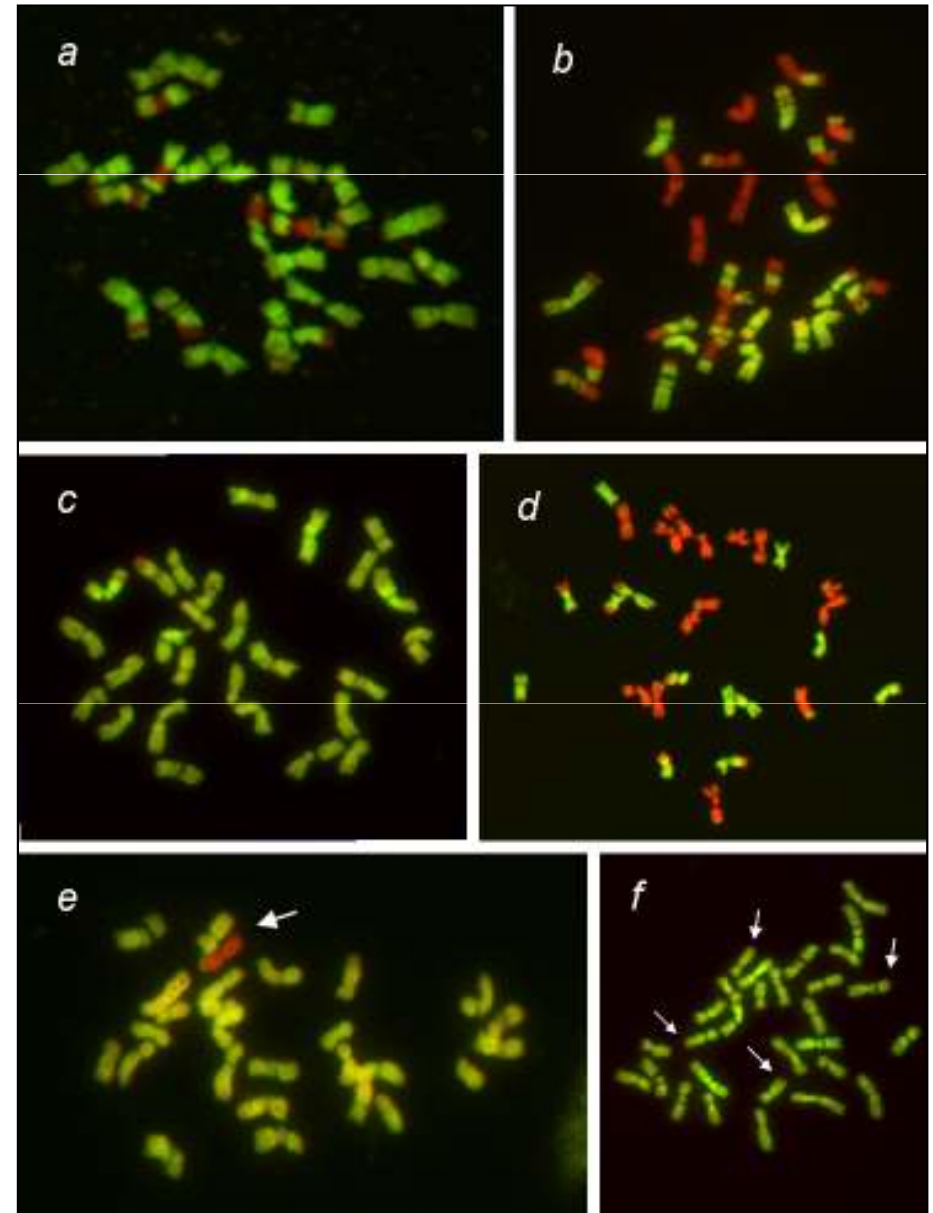
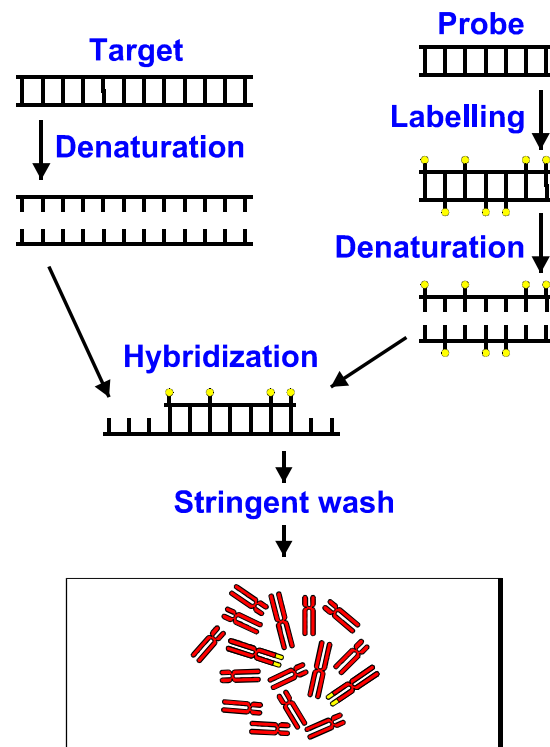
- Genetic diversity within species is limited – new genetic resources needed
- Interspecific and intergeneric crosses
- *Lolium x Festuca = xFestulolium*
- 1967 – Dr. Antonín Fojtík
- 1988 – registration of Felina

Questions

- What is the genomic constitution of our cultivars?
- Which method is the best to use?
- Is it possible to introgress any segments from any part of genome?

Genomic constitution of hybrid cultivars

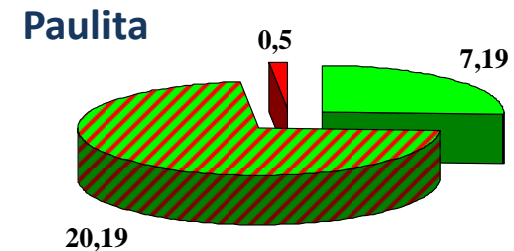
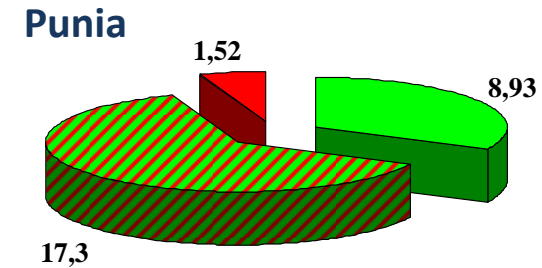
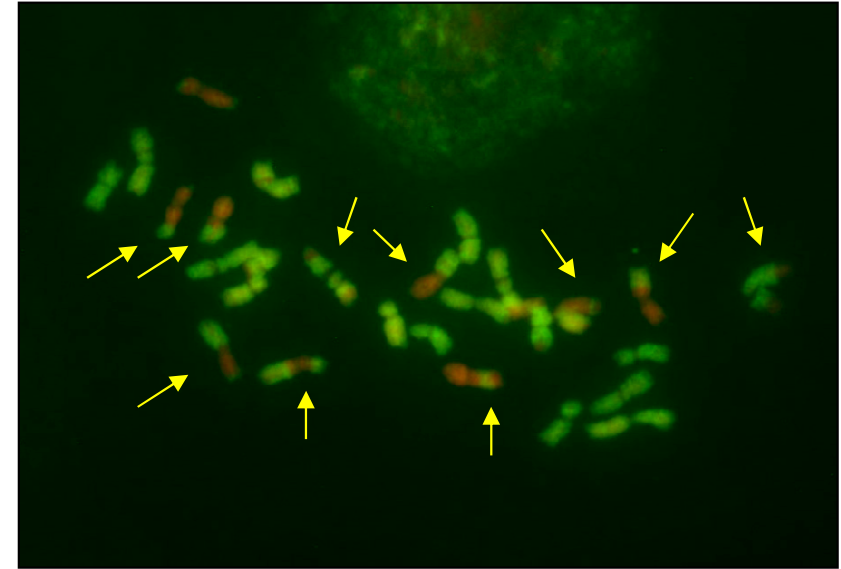
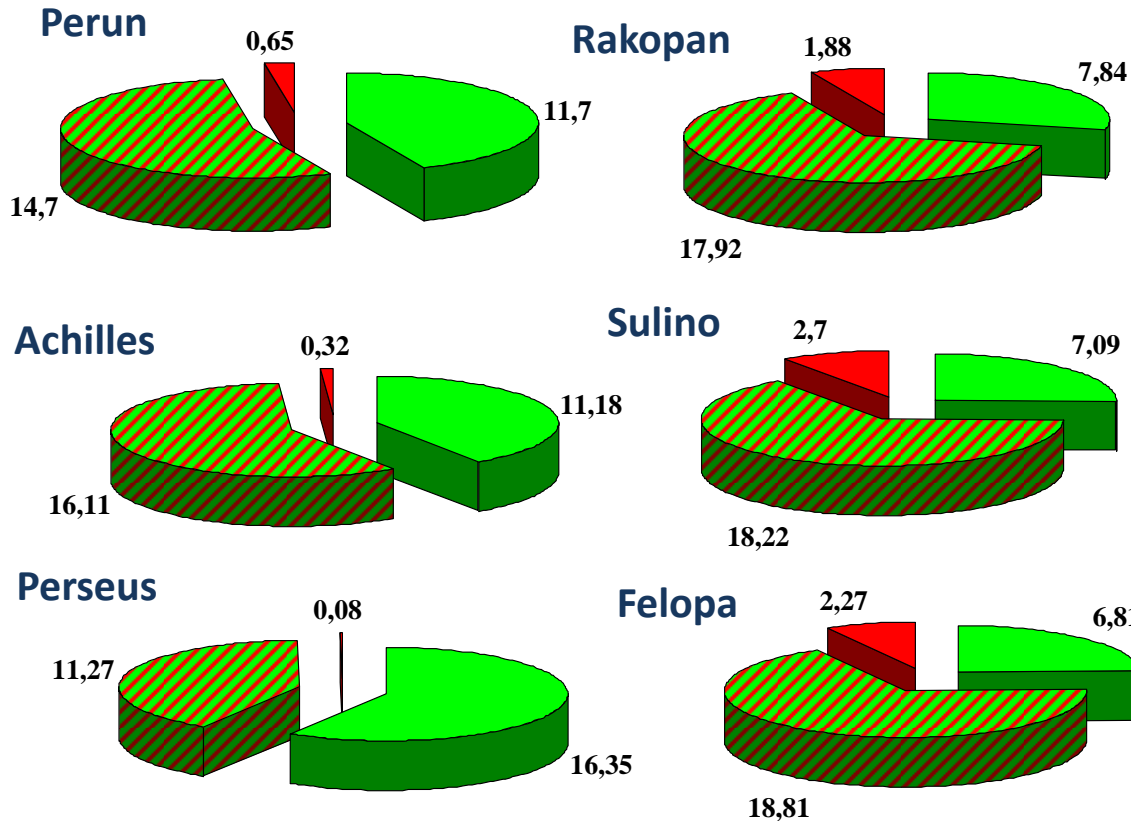
Fluorescent *in situ* hybridization (FISH)
Genomic *in situ* hybridization (GISH)



Number of Festulolium cultivars: ~42
Number of screened cultivars: 25

Genomic constitution of hybrid cultivars

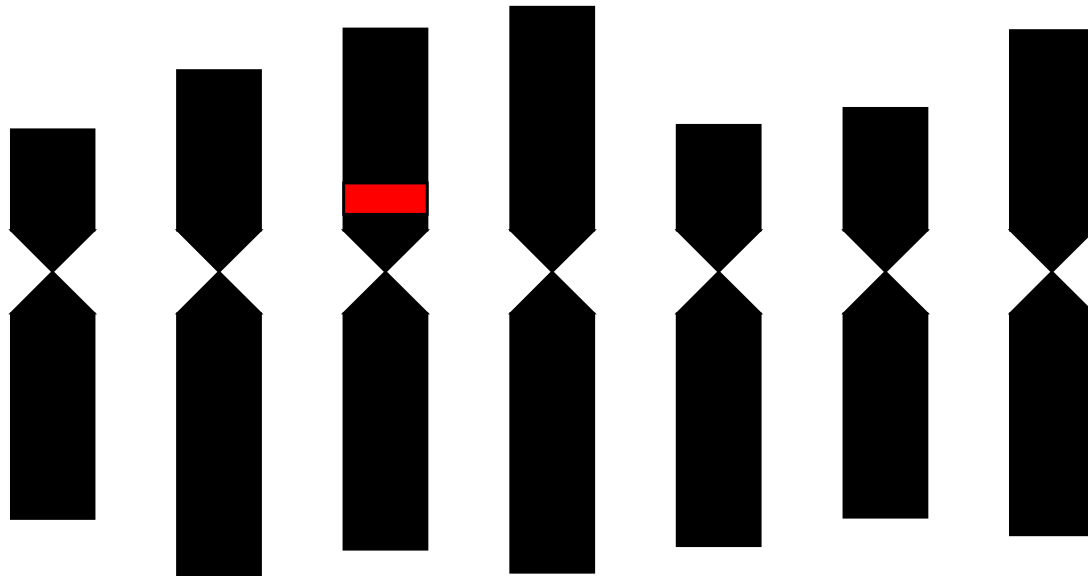
L. multiflorum *F. pratensis*



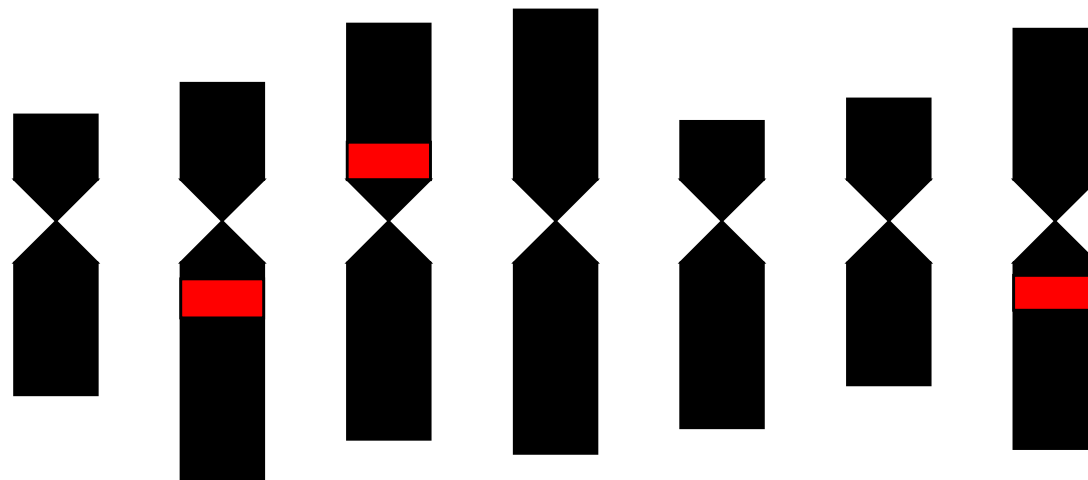
■ *L. multiflorum*
 ■ *F. pratensis*
 ■ Recombination LmxFp

*) number of chromosomes on average

Karyotyping

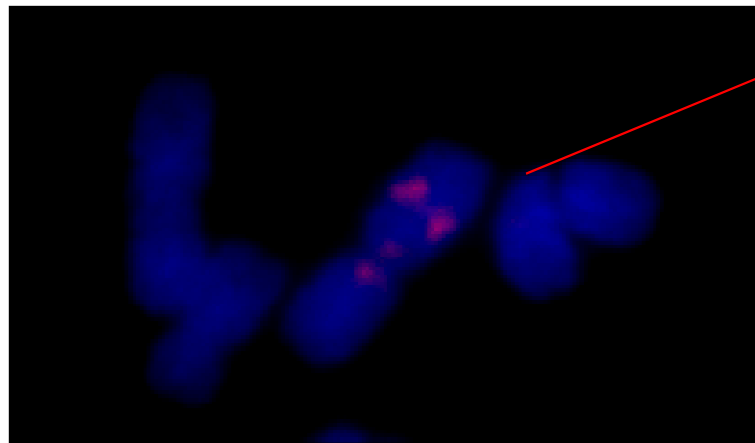
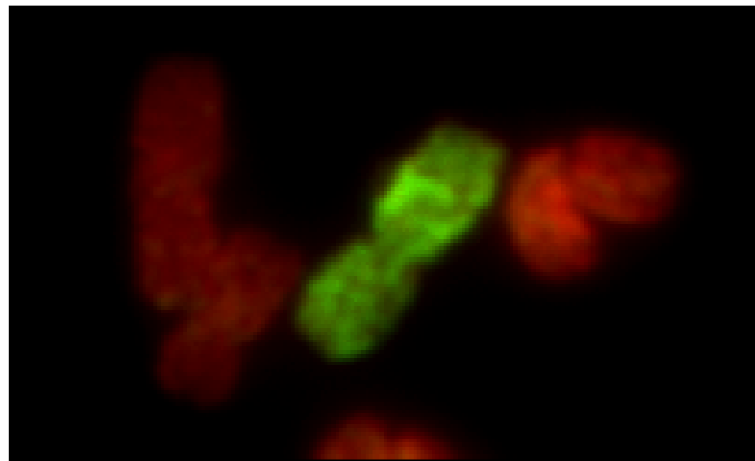


| 1 μ m *F. pratensis*

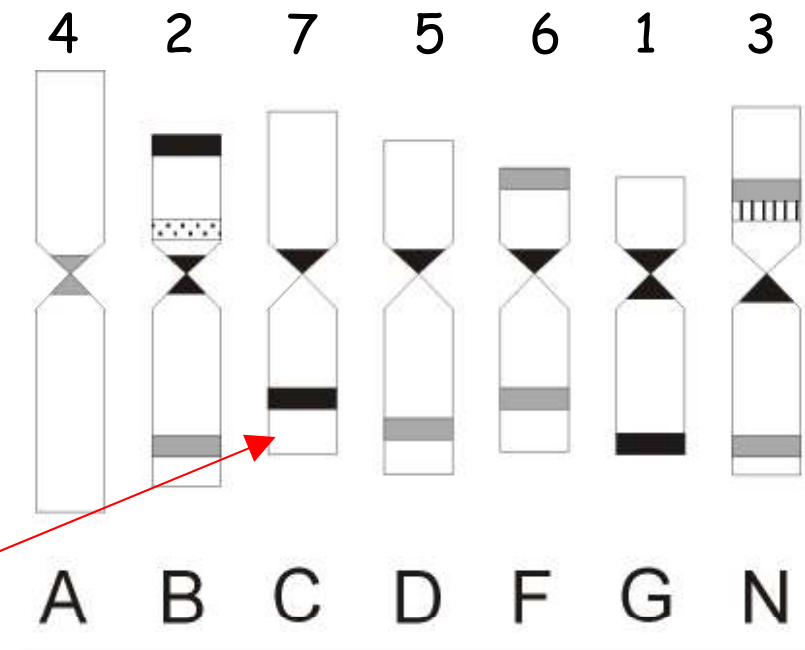


| 1 μ m *L. multiflorum*

Cytogenetic mapping in *F. pratensis*

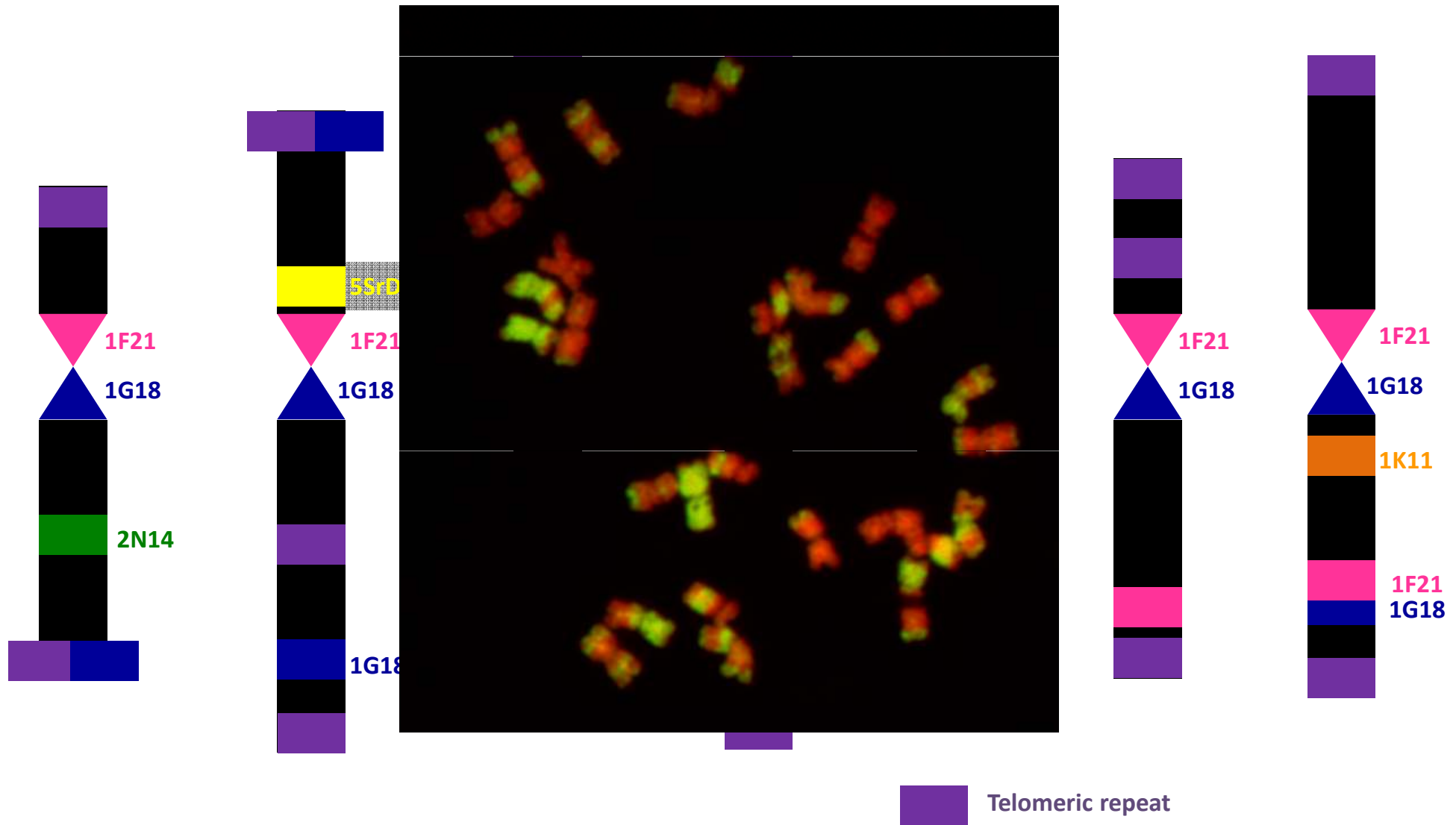


Chromosome C



- BAC 1G18
- BAC 1G18 (weak signal)
- ▤ 5S rDNA
- ▤ 45S rDNA

Usage of BAC-FISH/GISH



DArTFest

A first DArT array for the *Festuca-Lolium* complex

Development of array:

Material: 5 species

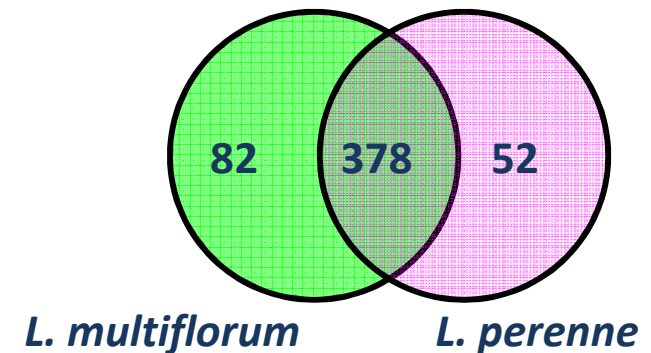
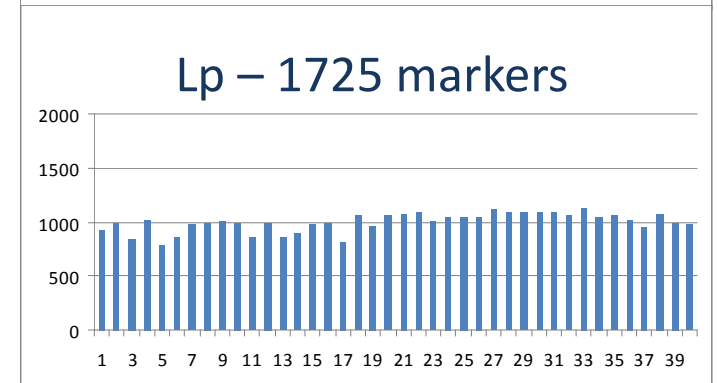
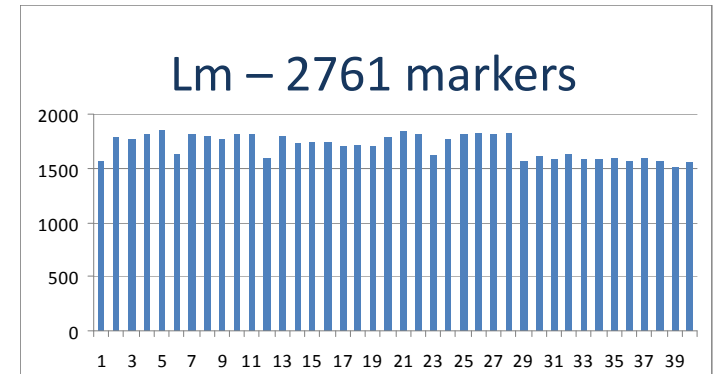
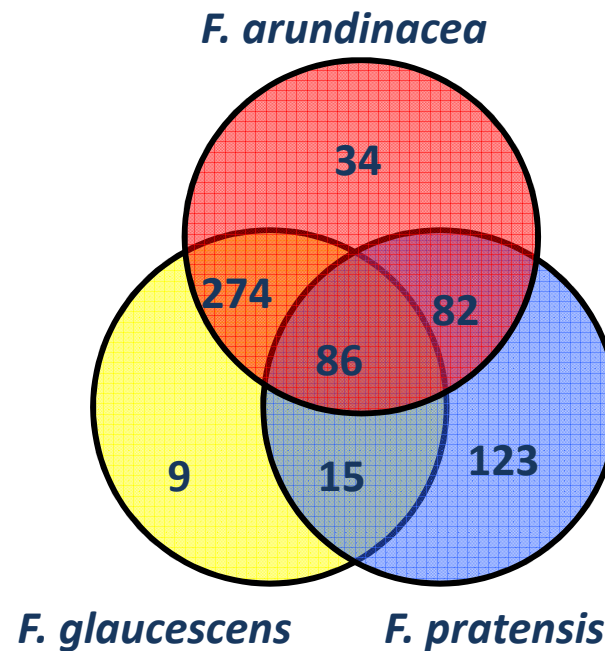
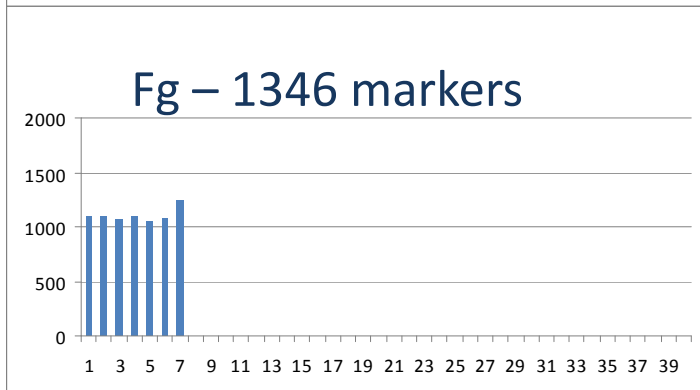
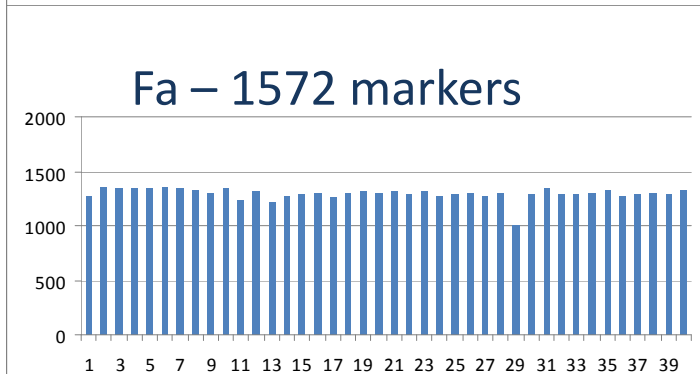
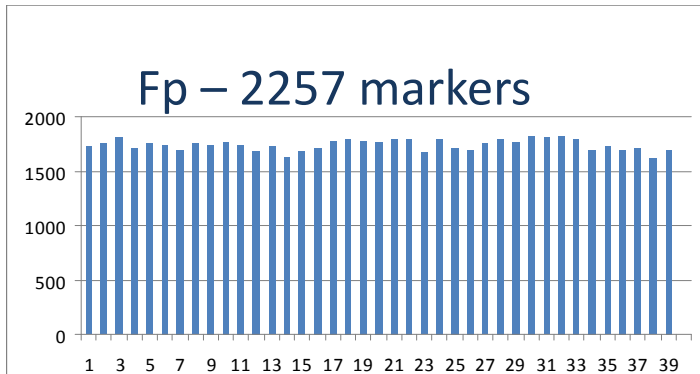
- 40 accessions of *Lolium perenne* (2x, 4x)
- 40 accessions of *L. multiflorum* (2x, 4x)
- 40 accessions of *Festuca pratensis* (2x, 4x)
- 40 accessions of *F. arundinacea* (6x)
- 7 accessions of *F. glaucescens* (4x)

Array: 7680 probes

Number of polymorphic markers: 3884

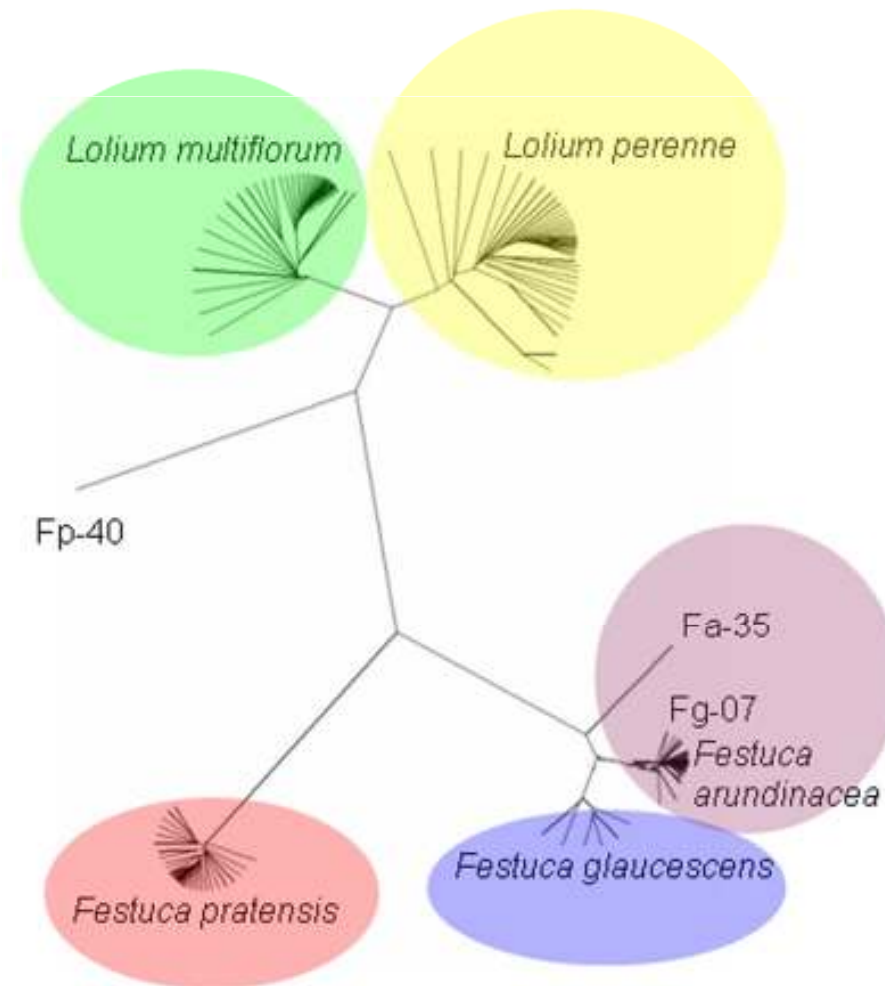
Current cost of analysis: Au\$5,000/94 samples

DArTFest : Genus- and species-specific markers



DArTFest : Genetic diversity

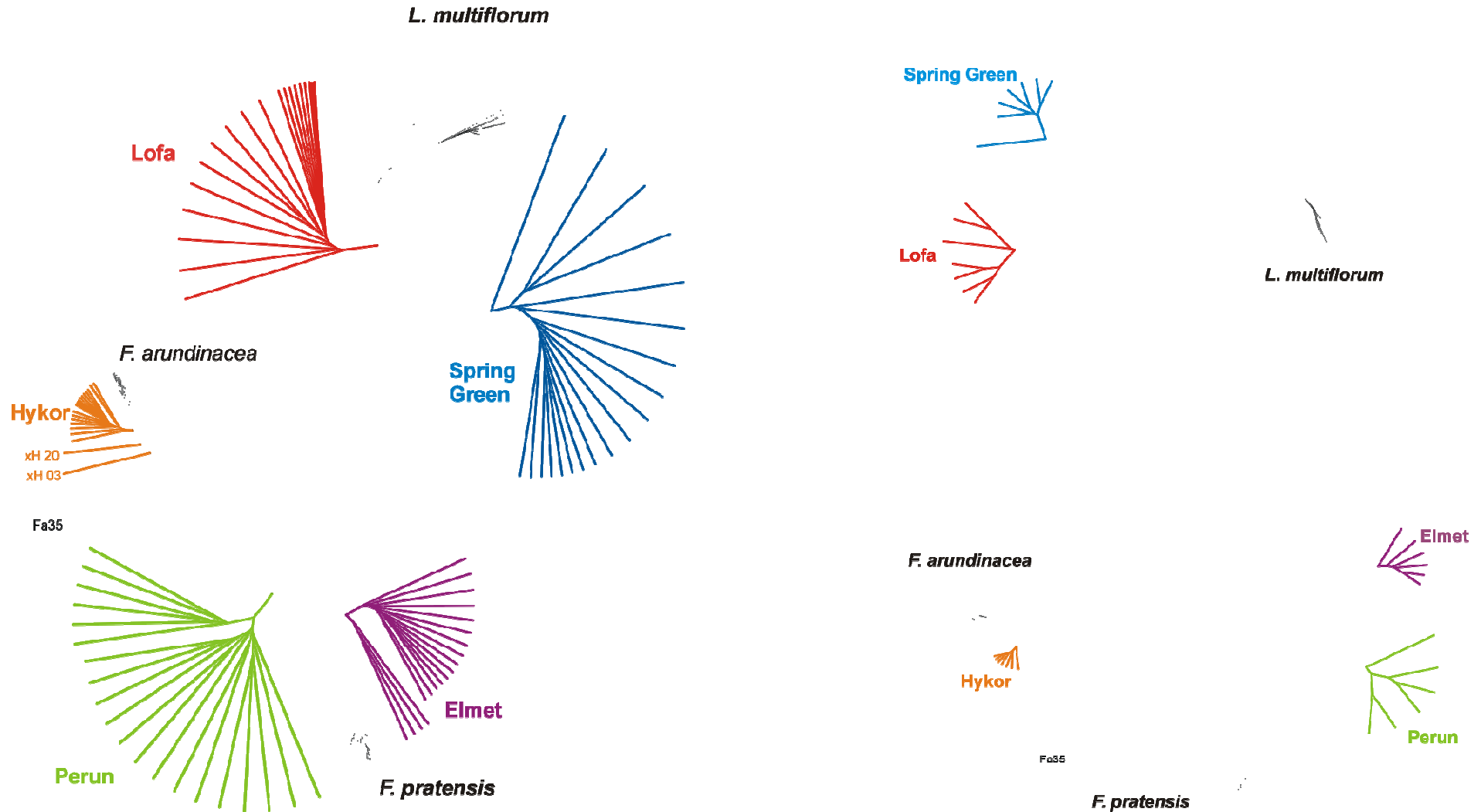
Based on the analysis of
2637 DArT markers



Diversity in hybrids

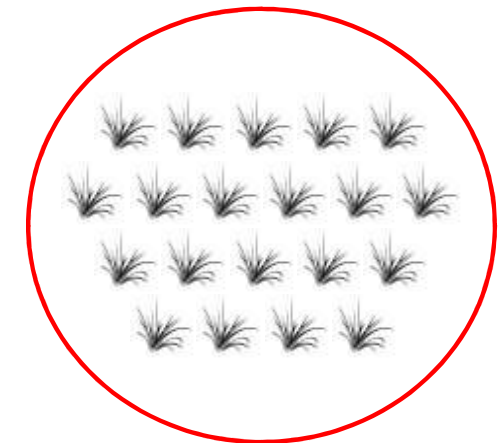
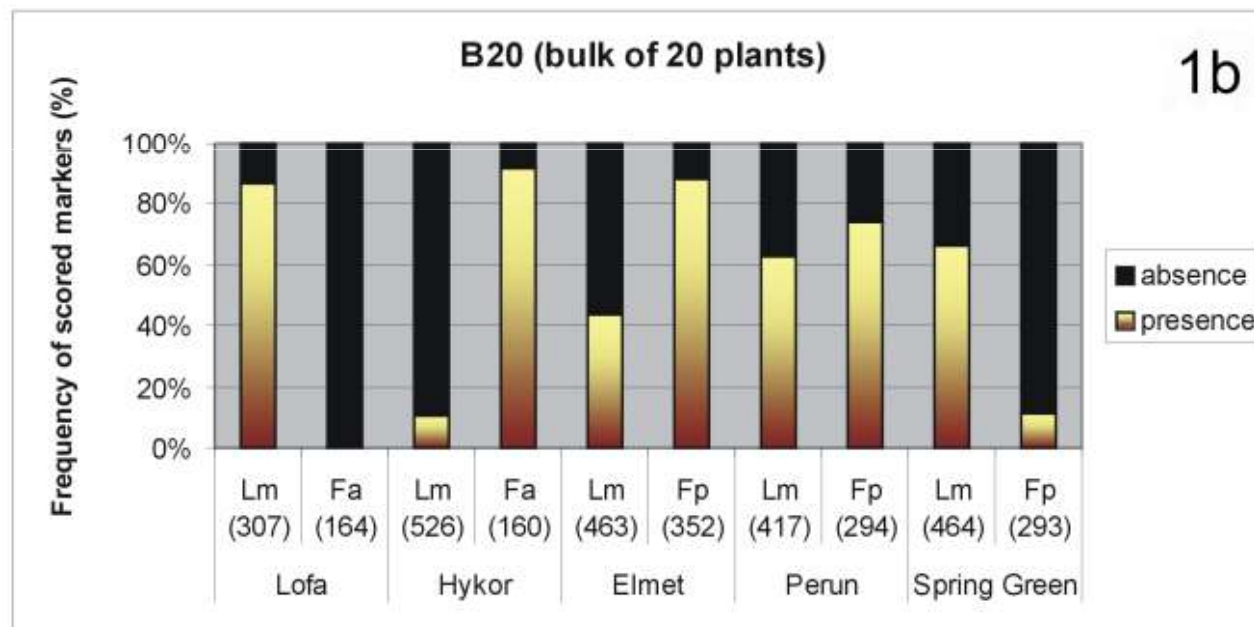
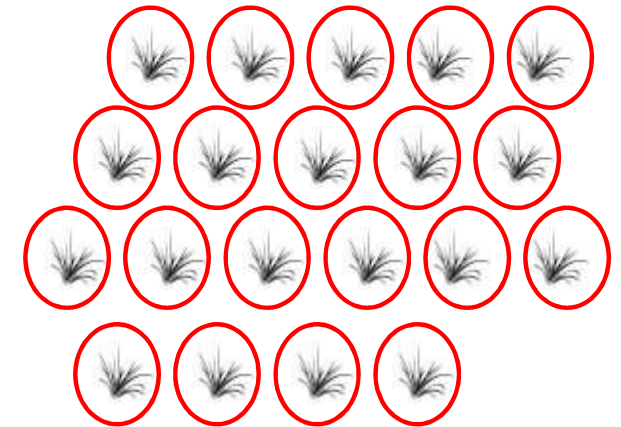
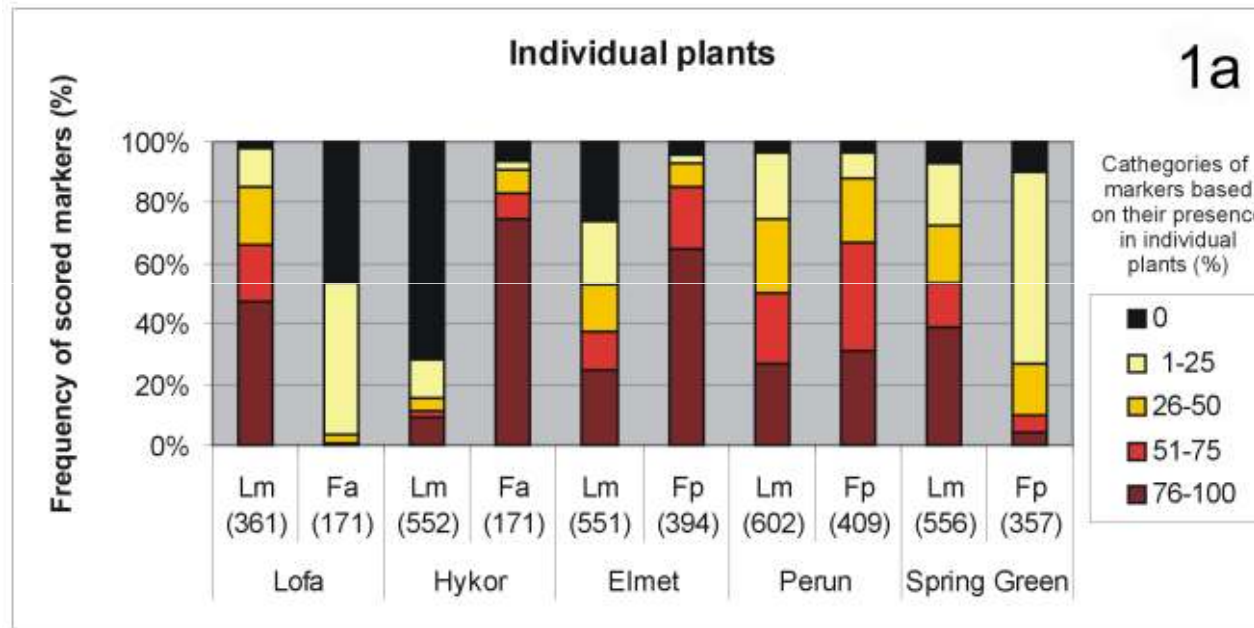
Individual plants

Bulked samples

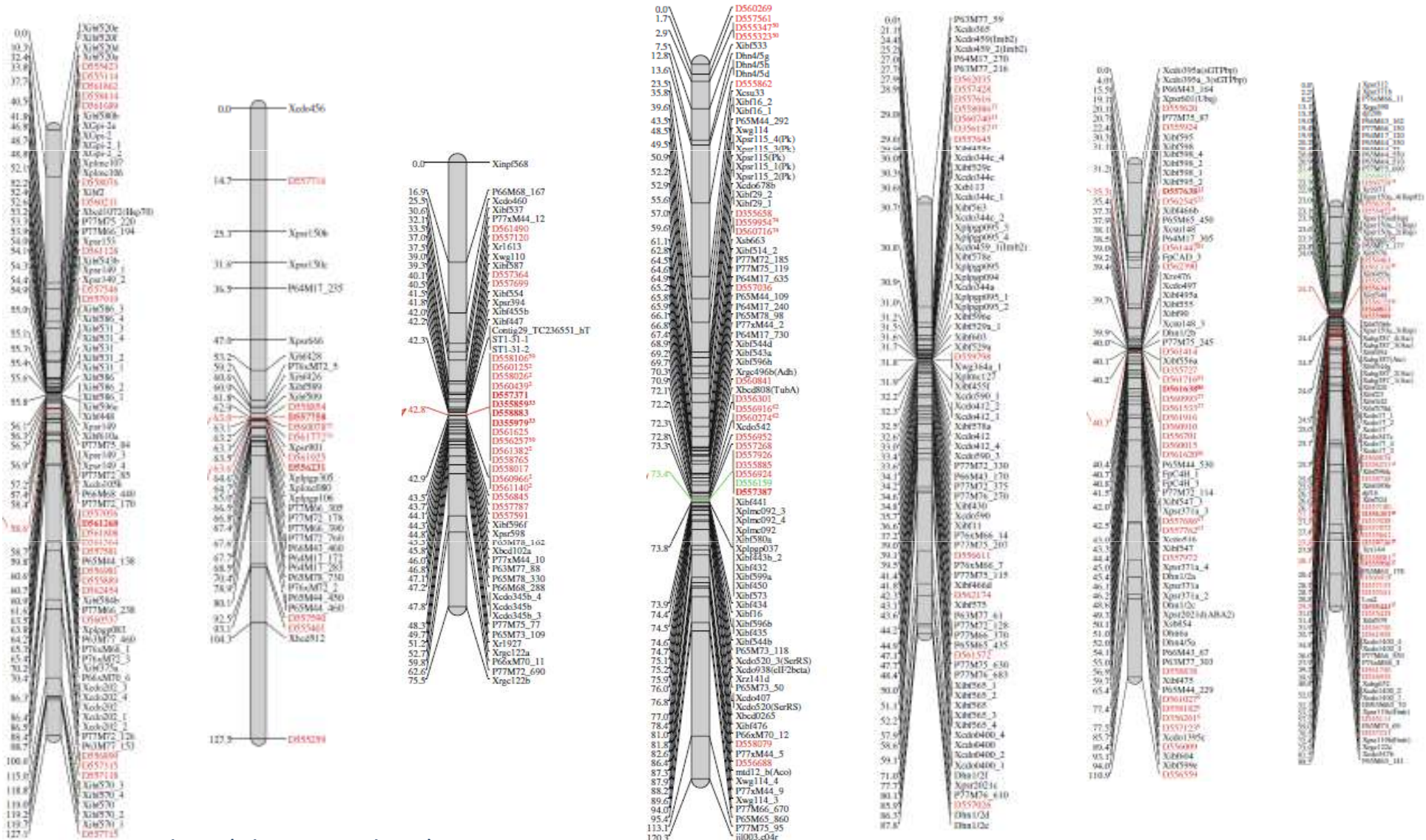


Based on 1471 informative DArT markers

Genomic constitution of hybrids

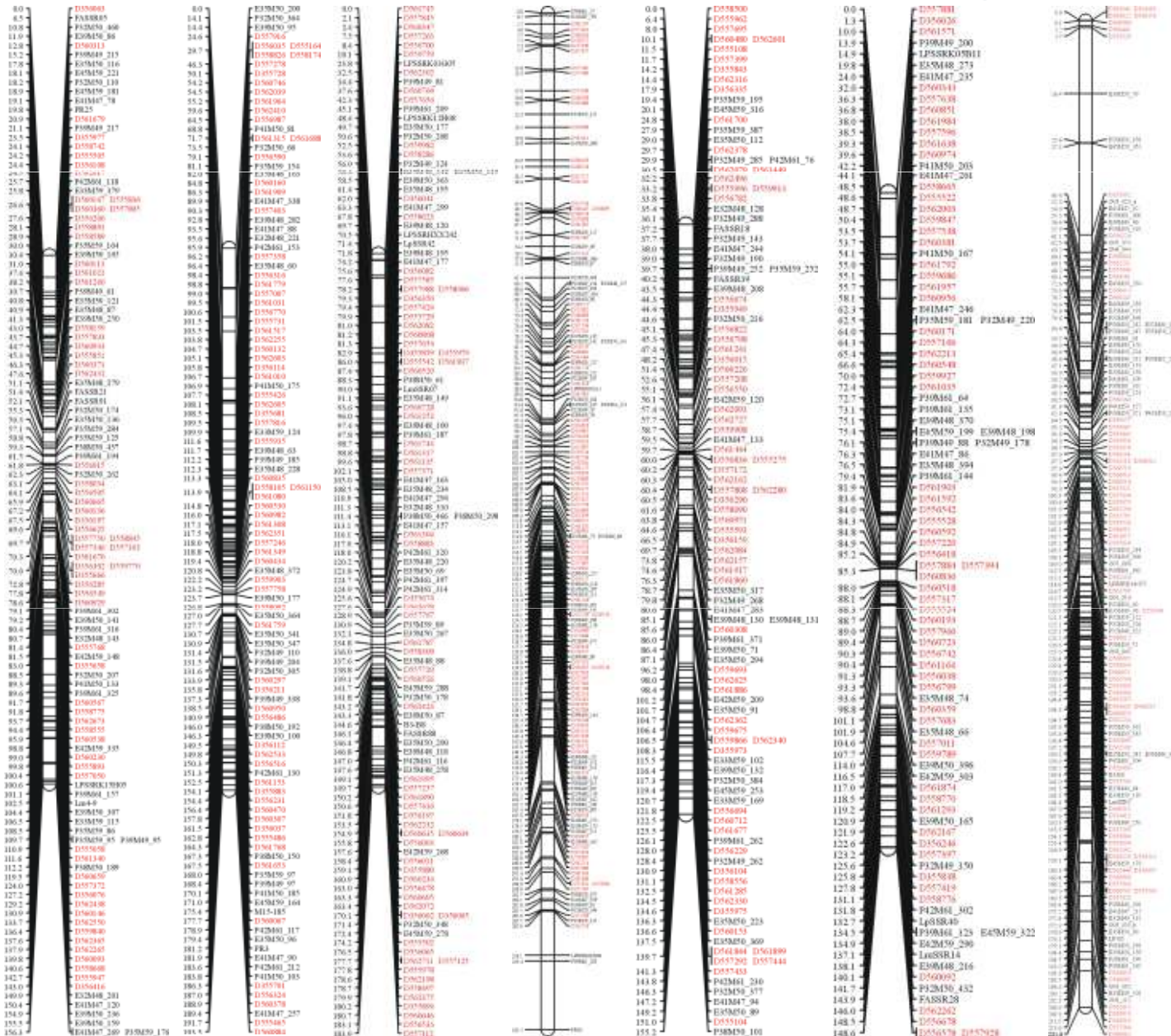


DArTFest : Genetic mapping (*Festuca pratensis*)



- 204 DArT markers (plus 466 others)
- The total length of genetic map 775cM (DArT genetic map by itself - 507cM; the original map - 605cM) = 1.16 cM/marker

DArTfest : Genetic mapping (*Lolium multiflorum*)



Total length: 1295 cM
 Number of DArT markers: 531
 Number of other markers: 352
 1.47 cM/marker

Sequencing

Up to date, we have sequenced over 1000 DArT markers

Results from 621 markers placed on genetic maps:

303,297 bp (=488.4 bp)

399 markers were singletons (64.3%)

222 markers were redundant (assigned to 90 marker bins)
= 489 non-redundant markers/bins

44 markers contained repetitive elements (=7.1%)

368 (59.3%) DArT markers with significant homology to expressed sequences

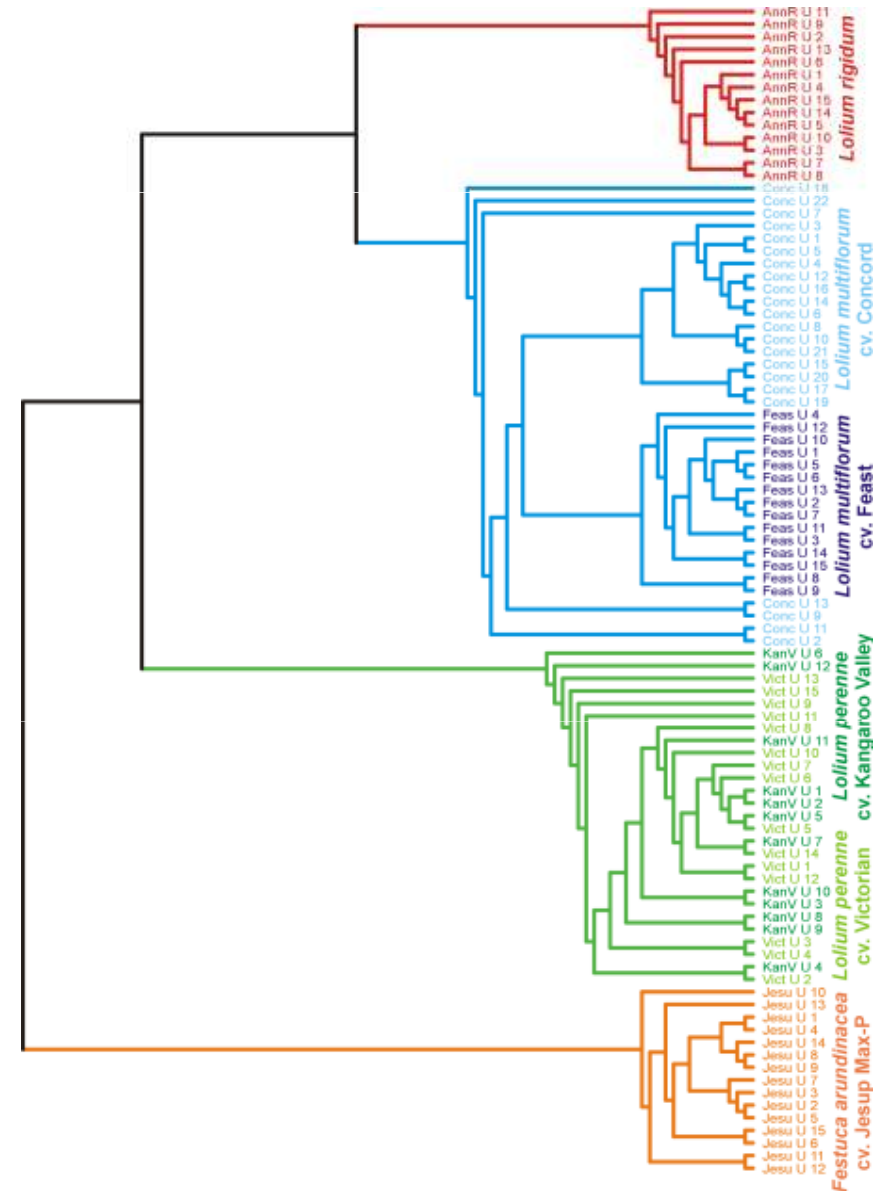
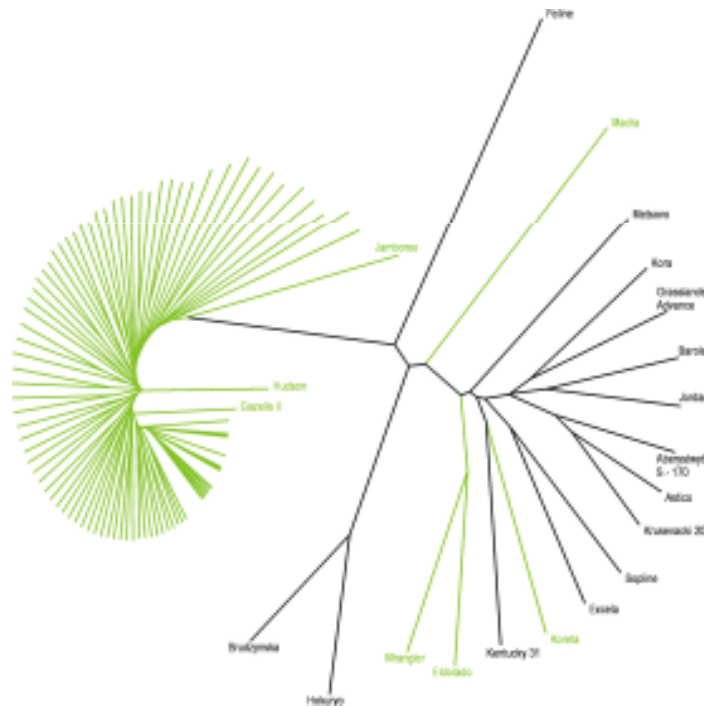
163 (26.2%) DArT markers with significant homology to known and hypothetical proteins

379 DArT markers (293 non-redundant bins) were identified as potentially gene-derived sequences



Current research using DArTFest

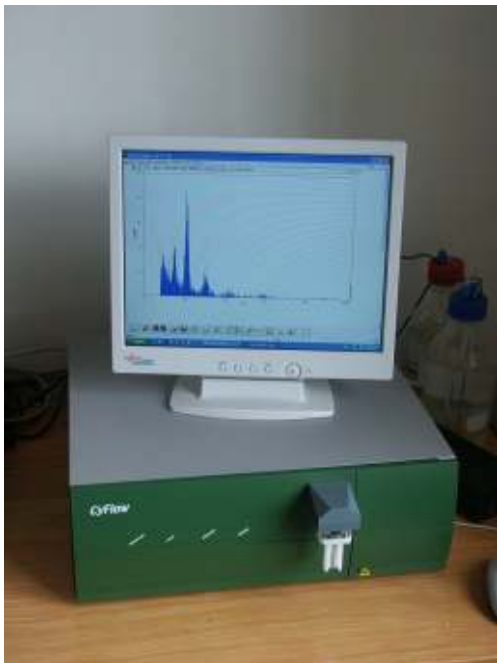
- Genetic diversity
- Genetic mapping
- Physical mapping
- Genetic constitution of hybrids
- Marker Assisted Selection



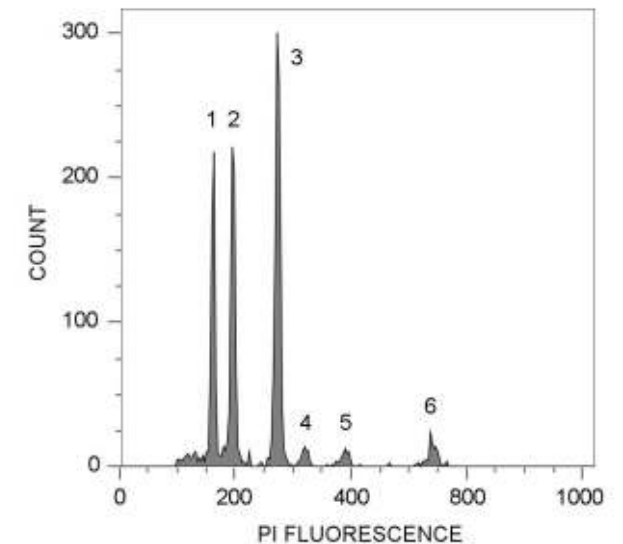
NGS - Next Generation Sequencing

Genome size estimation

Species	Cultivar	2C (pg)		1Cx (pg)	1C (Mbp)	Mean CV values	Ploidy level	No. of chromosomes
		Mean	SD					
<i>Festuca pratensis</i>	Kolumbus	6,49	0,078	3,25	3175	2,96	2	14
<i>Festuca pratensis</i>	Patra	13,01	0,050	3,25	5150	2,41	4	28
<i>Festuca arundinacea</i>	Kora	17,45	0,078	2,91	8534	2,80	6	42
<i>Lolium perenne</i>	Handicap	5,36	0,047	2,68	2623	2,16	2	14
<i>Lolium perenne</i>	Korok	11,19	0,057	2,80	5470	2,27	4	28
<i>Lolium multiflorum</i>	Prolog	5,25	0,042	2,62	2567	2,35	2	14
<i>Lolium multiflorum</i>	Lubina	10,99	0,092	2,75	5374	1,96	4	28



CyFlow flow cytometer (Partec GmbH., Münster, Germany)



1. *Lolium multiflorum* (2x)
2. *Festuca pratensis* (2x)
3. *Pisum sativum* (2x)

Molecular size of chromosomes (arms)

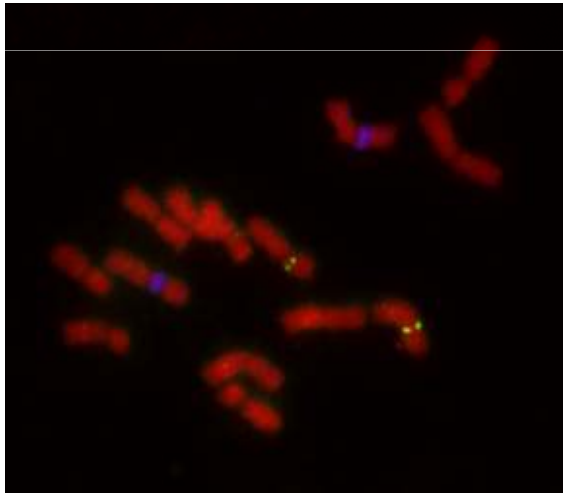
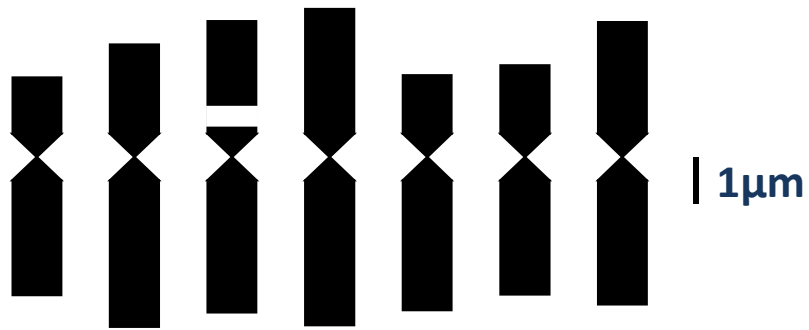
Lolium multiflorum Lam. (2n=4x=28)

	Length (μm)	%	1C (Mbp)		Length (μm)	%	1C (Mbp)		Length (μm)	%	1C (Mbp)	Cen-Index
1	3,37	11,01	296	1S	1,28	4,18	112	1L	2,09	6,83	183	0,38
2	4,77	15,58	419	2S	1,66	5,42	146	2L	3,11	10,16	273	0,35
3	4,99	16,30	438	3S	2,37	7,74	208	3L	2,62	8,56	230	0,47
4	5,33	17,41	468	4S	2,54	8,30	223	4L	2,79	9,11	245	0,48
5	3,68	12,02	323	5S	1,20	3,92	105	5L	2,48	8,10	218	0,33
6	3,44	11,23	302	6S	1,47	4,80	129	6L	1,97	6,43	173	0,43
7	5,04	16,46	442	7S	2,31	7,54	203	7L	2,73	8,92	240	0,46

Festuca pratensis Huds. (2n=2x=14)

	Length (μm)	%	1C (Mbp)		Length (μm)	%	1C (Mbp)		Length (μm)	%	1C (Mbp)	Cen-Index
1	4,67	11,74	373	1S	1,71	4,30	137	1L	2,96	7,44	237	0,37
2	6,07	15,24	485	2S	2,42	6,09	194	2L	3,64	9,15	291	0,40
3	6,25	15,70	499	3S	2,92	7,34	233	3L	3,33	8,36	266	0,47
4	6,79	17,07	543	4S	3,18	7,99	254	4L	3,61	9,08	289	0,47
5	5,04	12,68	403	5S	1,76	4,41	140	5L	3,29	8,27	263	0,35
6	4,93	12,38	394	6S	1,97	4,95	158	6L	2,95	7,42	236	0,40
7	6,05	15,21	484	7S	2,90	7,28	231	7L	3,16	7,93	252	0,48

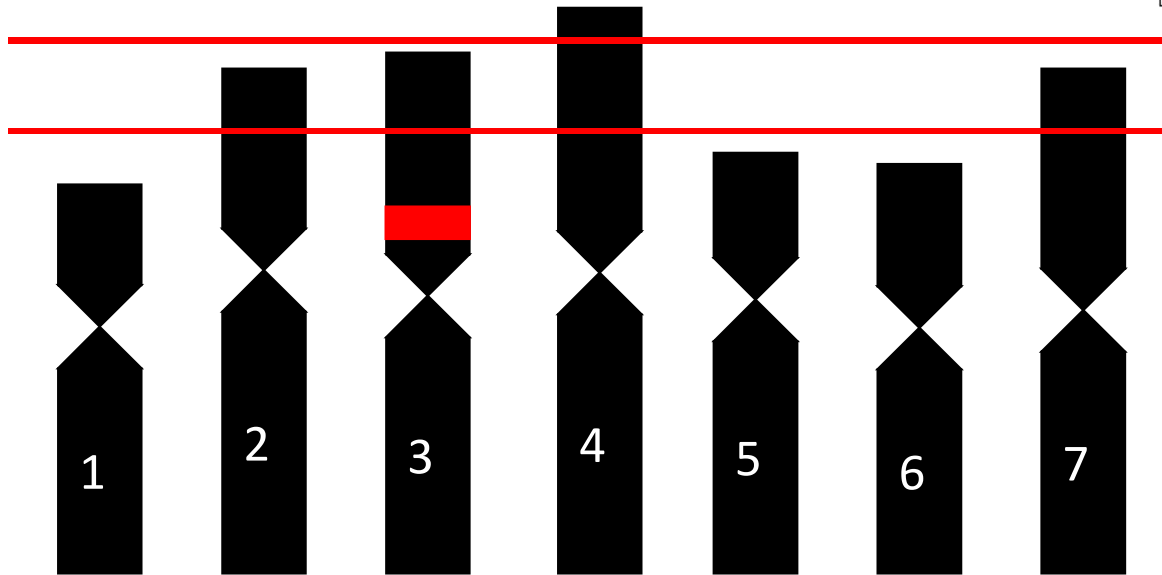
meadow fescue



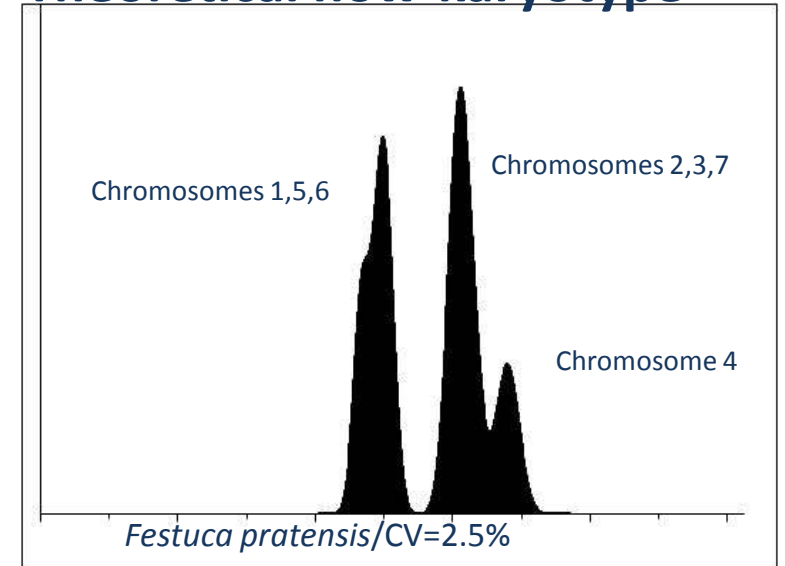
KARYOTYPING OF MEADOW FESCUE

Chromosome	1	2	3	4	5	6	7
Size (Mbp)	373	485	499	543	403	394	484
Cen							
Tel							
45SrDNA							
5SrDNA							
1F21							
1G18							
2N14							
2O16							
1K11							

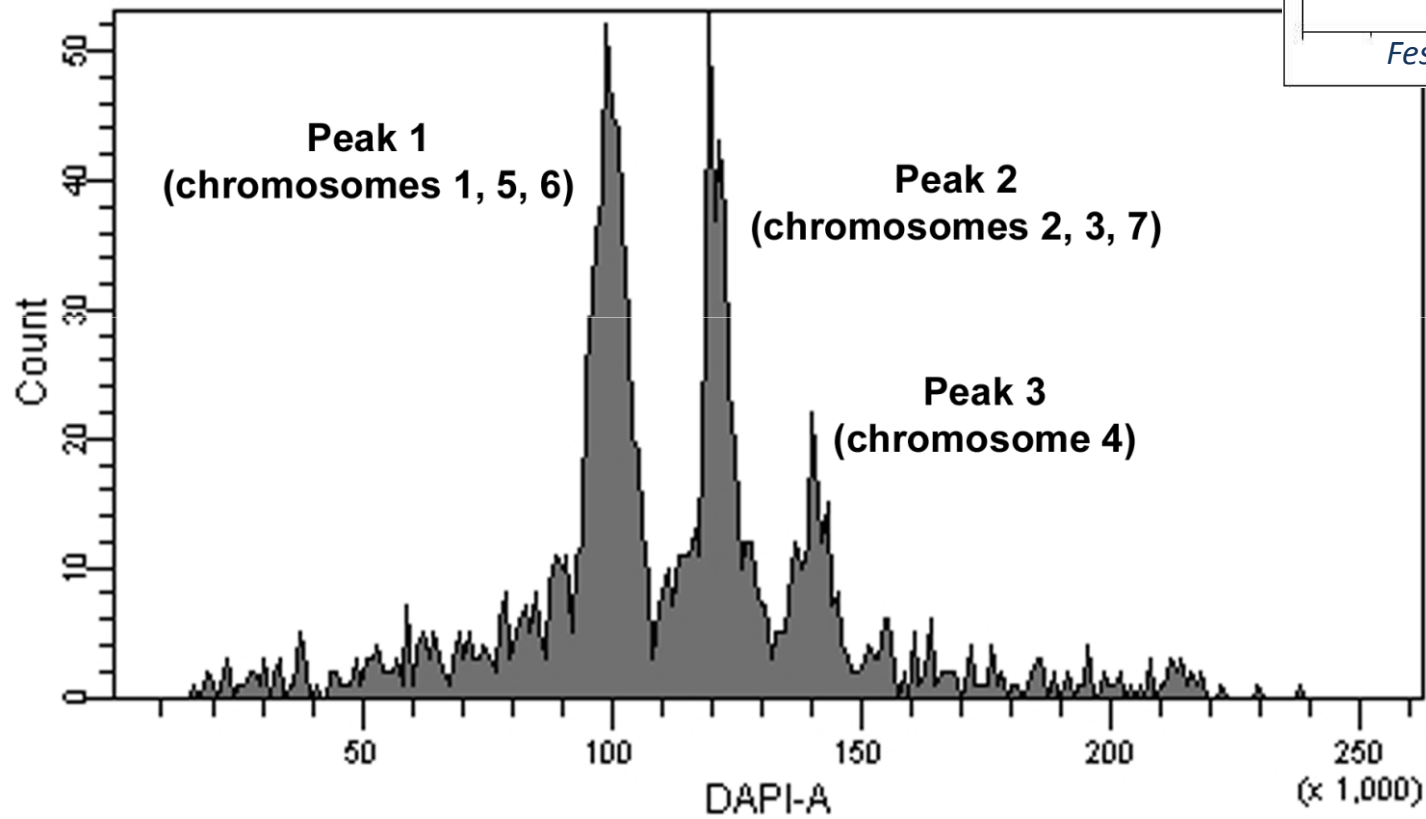
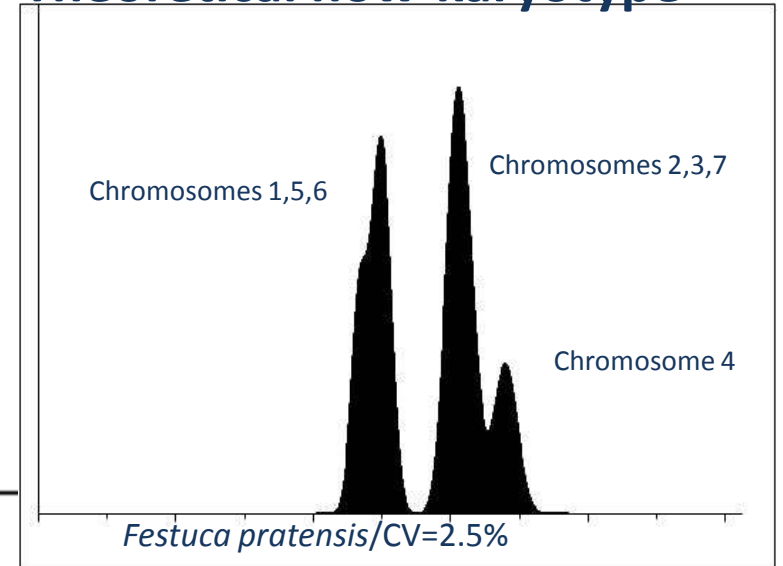
Sorting individual/groups of chromosomes



Theoretical flow-karyotype

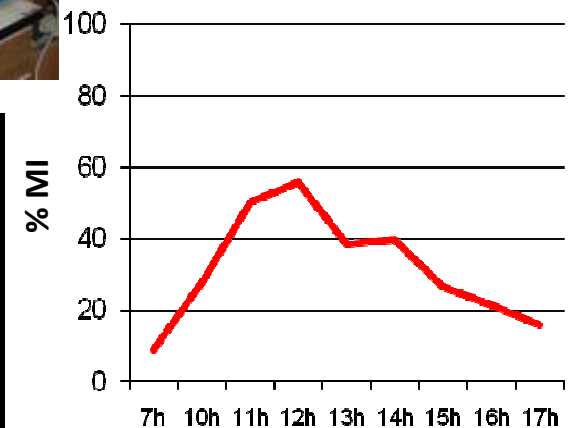
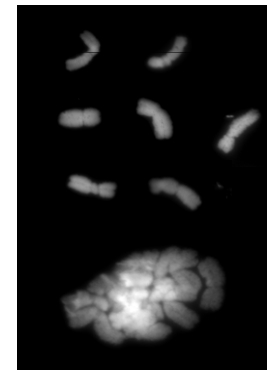
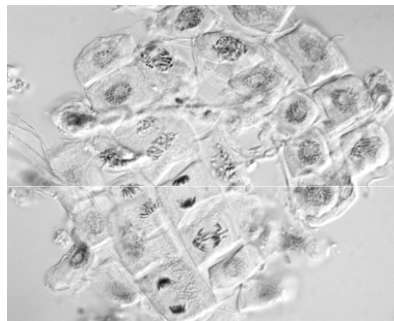
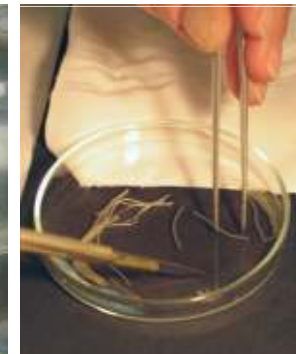
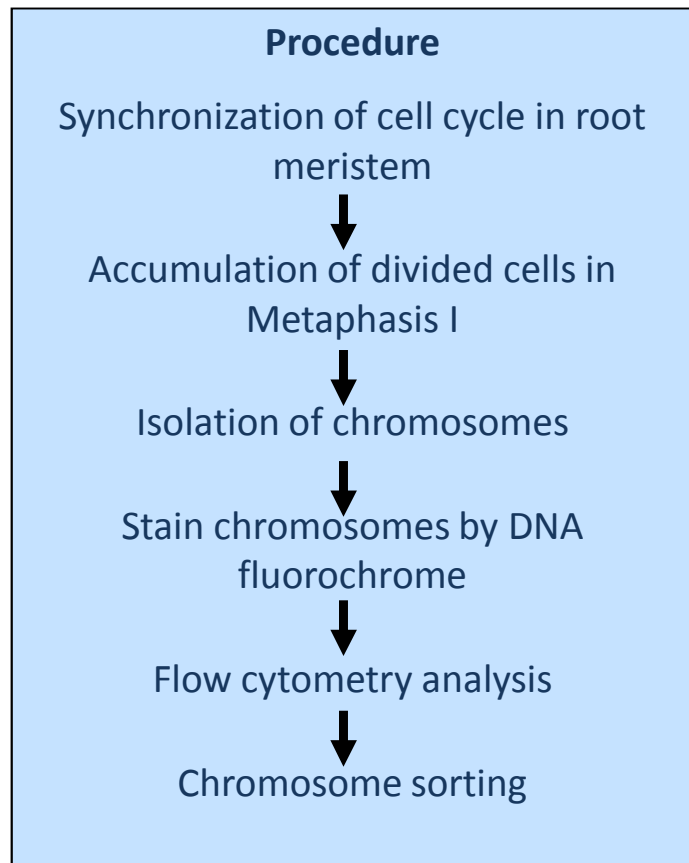


Theoretical flow-karyotype



synchronization of cell cycle and chromosome sorting

sample for sorting is a suspension of intact mitotic chromosomes



18h
1.5mM HU

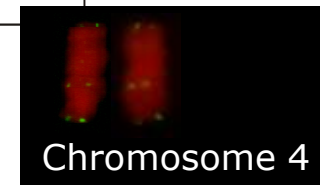
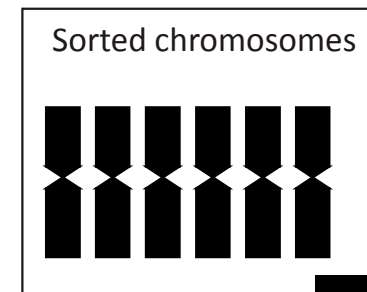
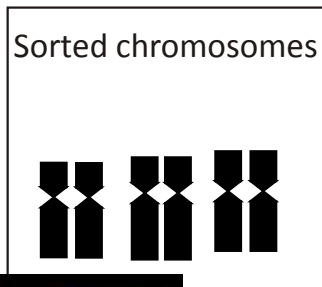
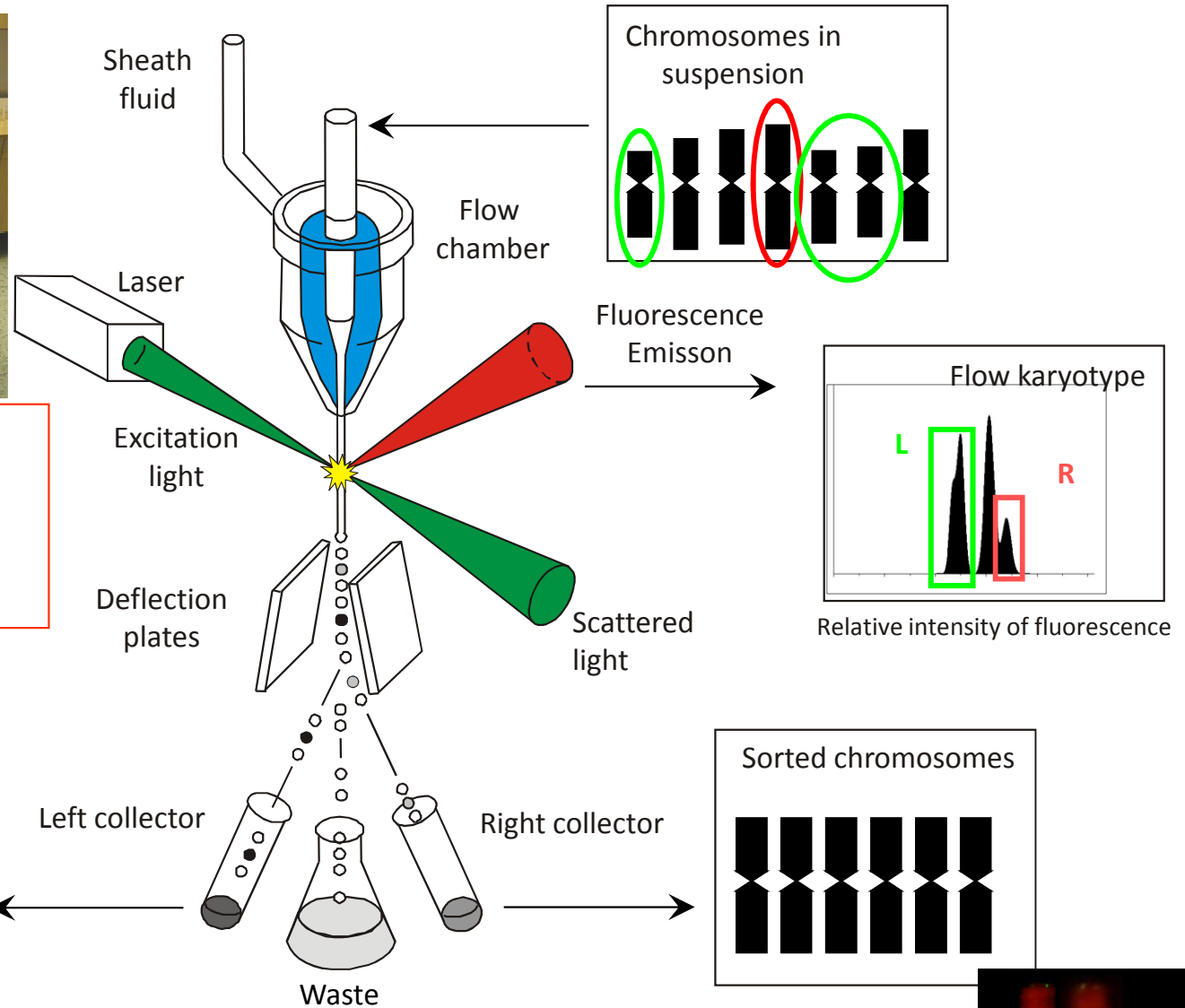
4.5h
Recovery time

2h
5uM oryzalin

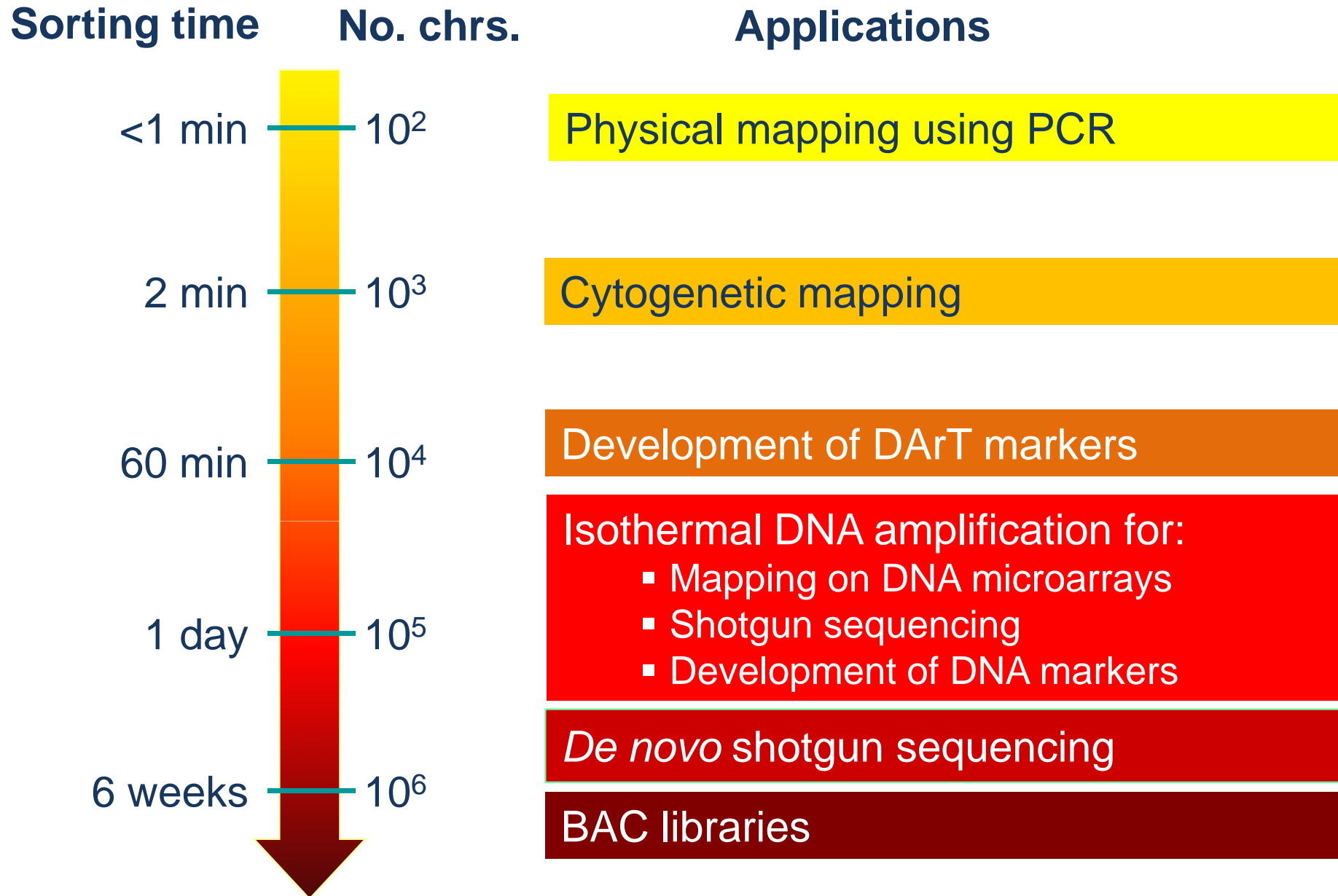
synchronization of cell cycle and chromosome sorting



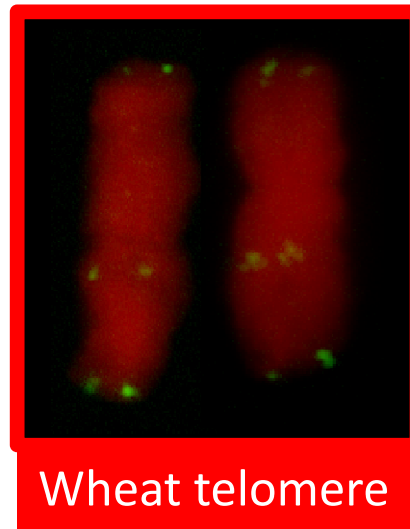
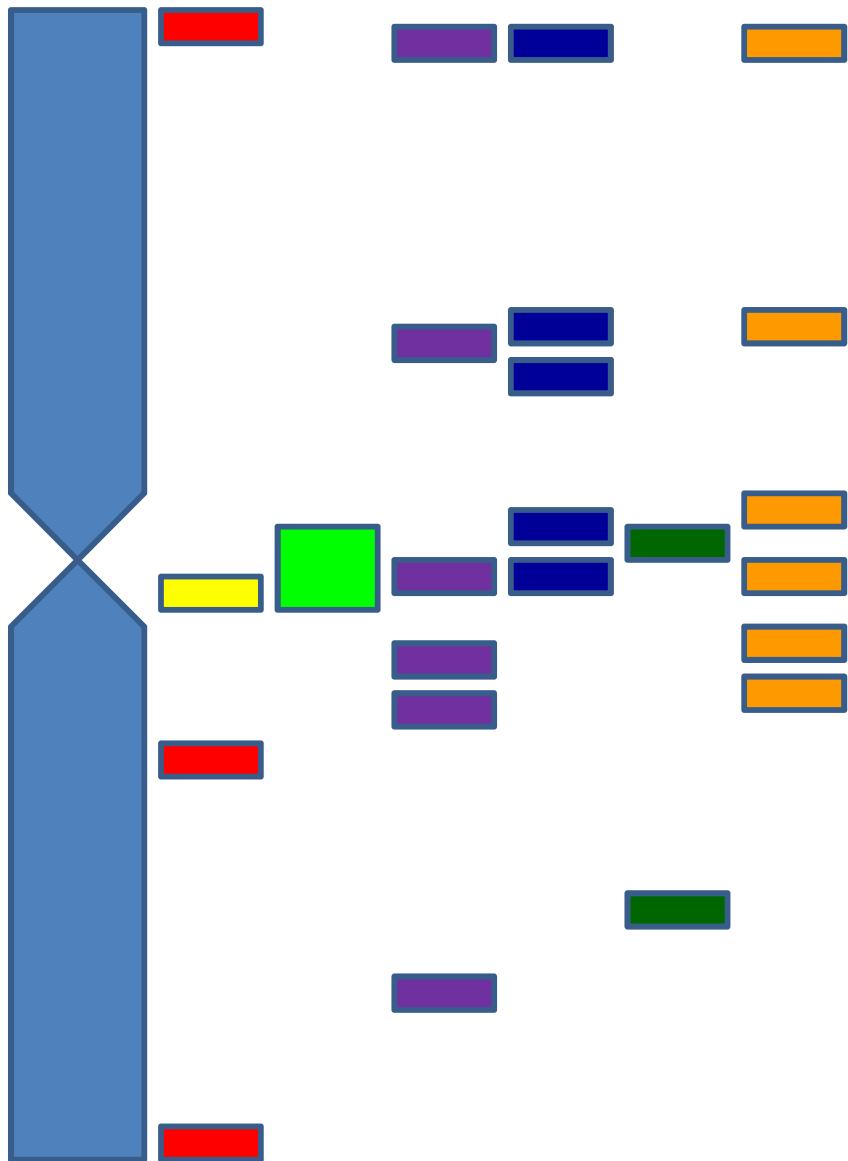
- Speed:
~1000 chromosomes/sec
- Yield per day:
2 - 5x10⁵ chromozómů



what to do with sorted chromosomes?



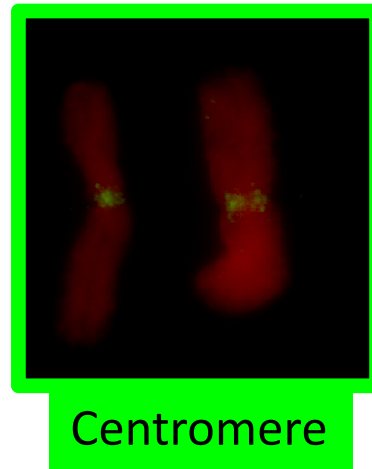
cytogenetics of chromosome 4F



Wheat telomere



BAC 1F21



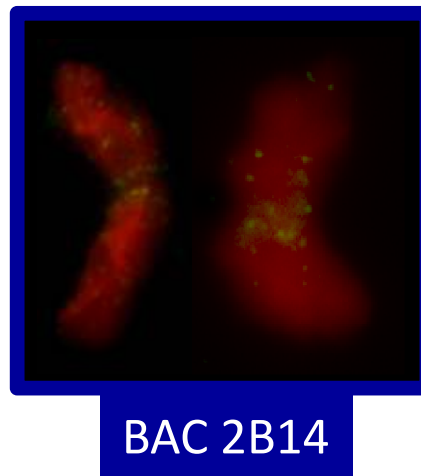
Centromere



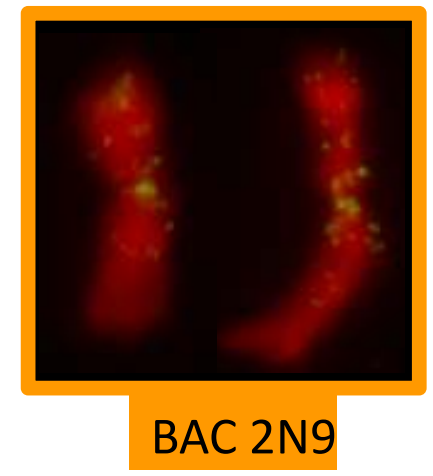
BAC 1G18



BAC 2D4



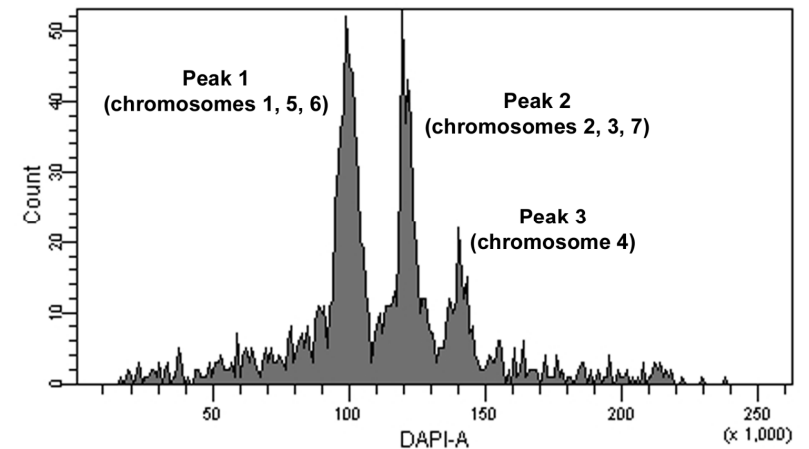
BAC 2B14



BAC 2N9

sequencing of chromosome 4F

- purity of sorted chromosome: 92.4%
- 140,000 copies of chromosome 4 amplified using GenomiPhi DNA Amplification Kit (GE Healthcare)
- sequenced by Illumina HiSeq2000 (IGA)
- 28,270 Mbp (over 50x coverage)



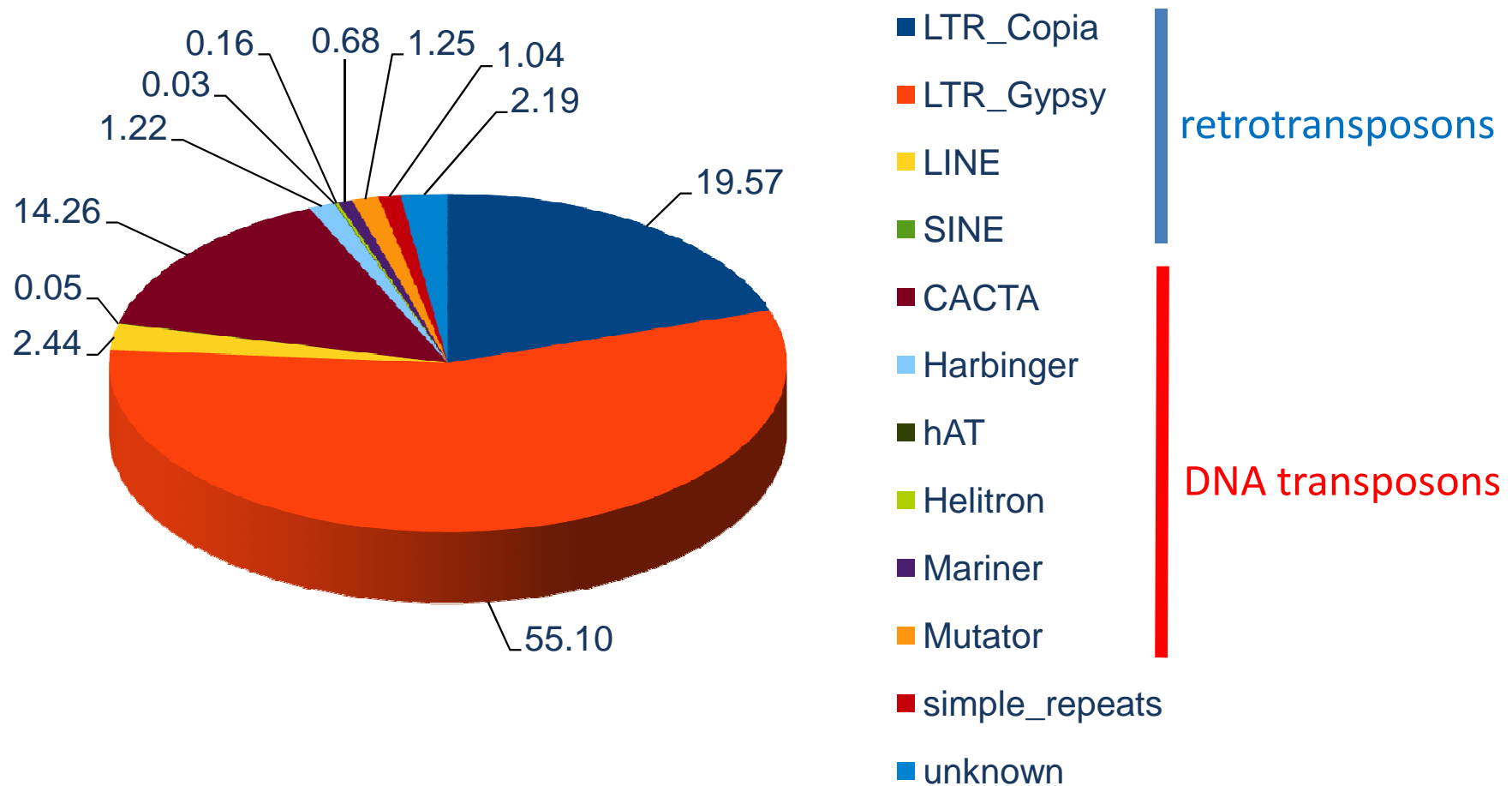
assembly

program SOAPdenovo with k-mer size 70

	4F
contigs	4,057,786
min bp	72
max bp	15,919
average	125.5
L50	132
# bp	509,357,265
GC-content [%]	42.58

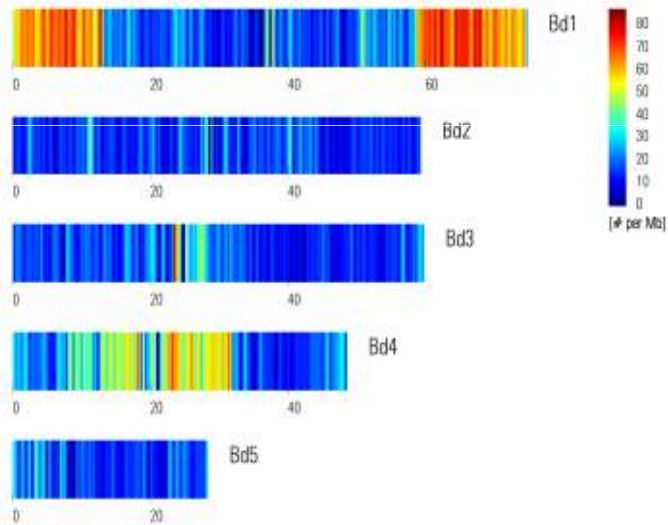
repeats

only **22.05%** sequences with hits in the repeat database

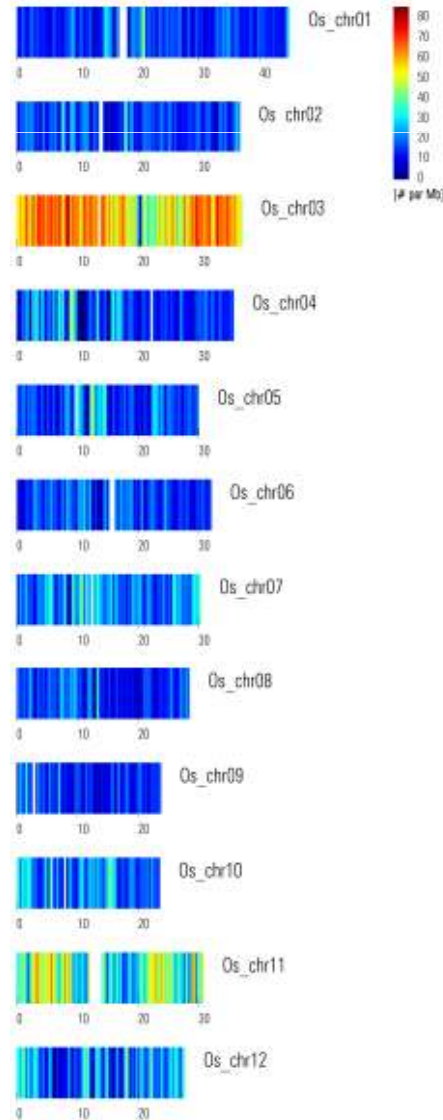


collinearity with brachy, rice and sorghum

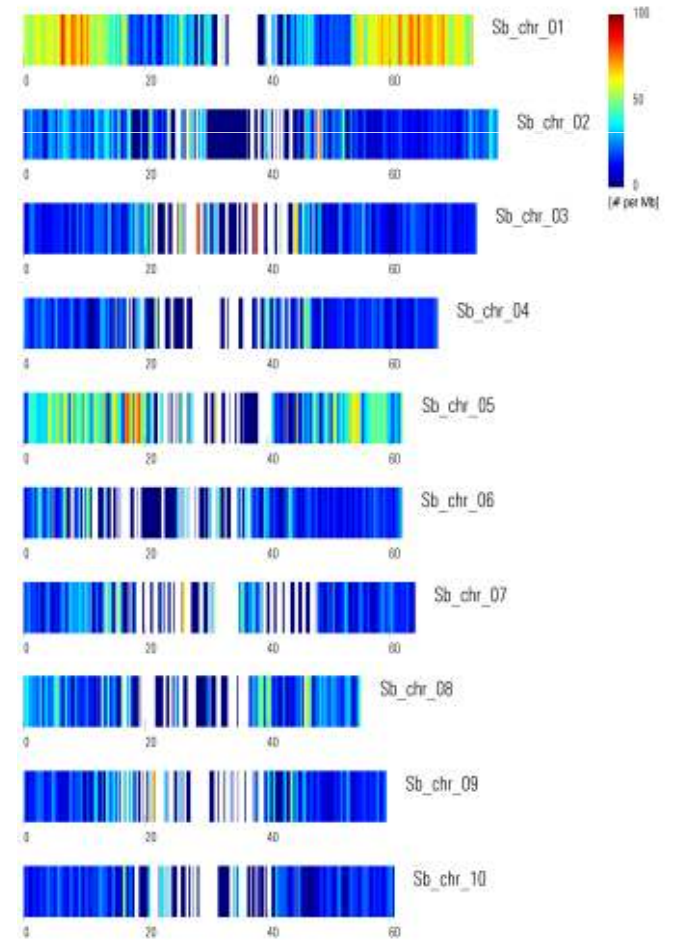
festuca_4f_vs_brachy



festuca_4f_vs_rice

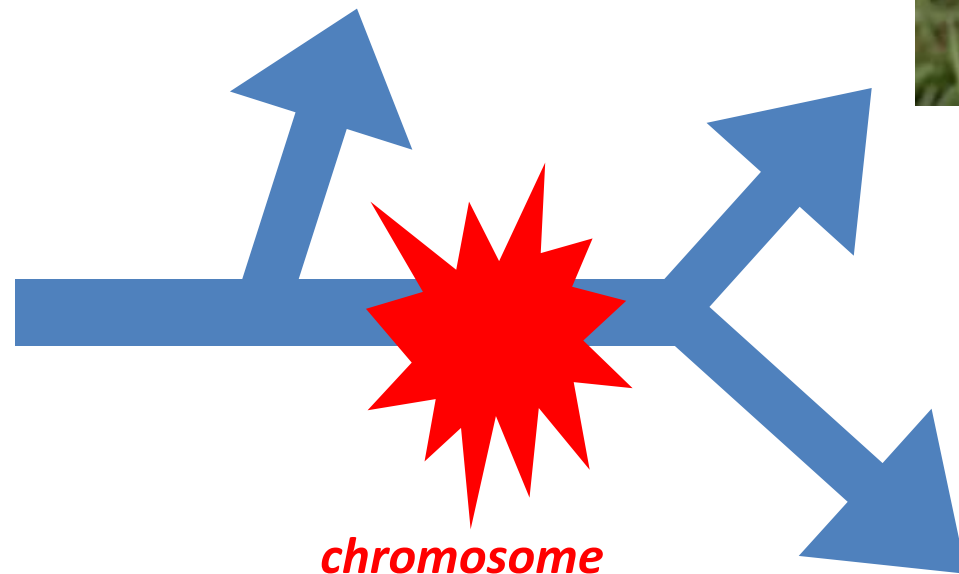
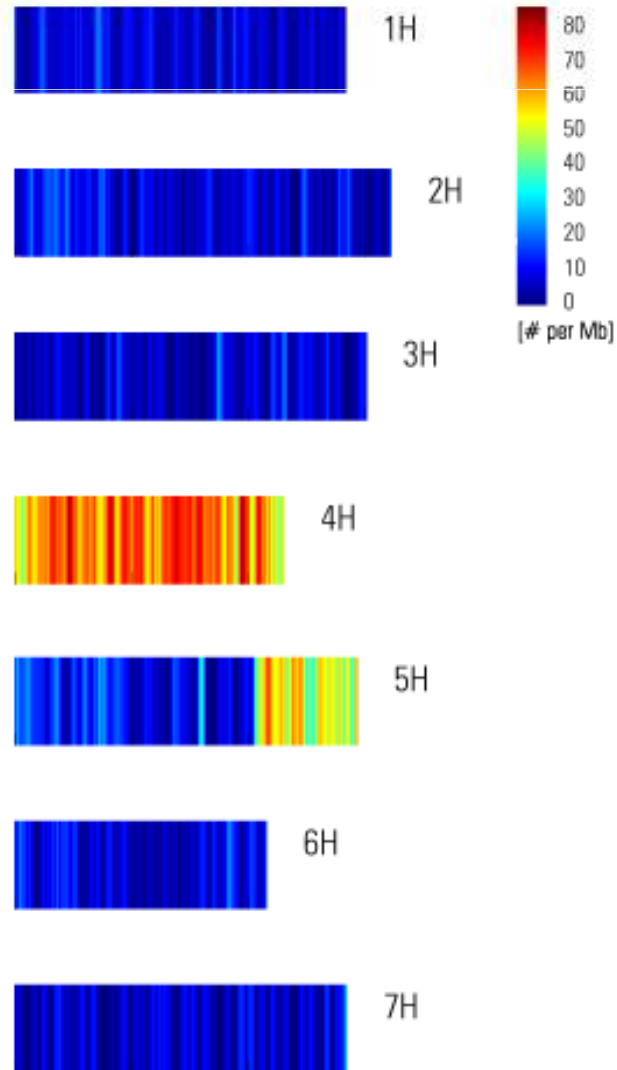


festuca_4f_vs_sorghum



collinearity with barley

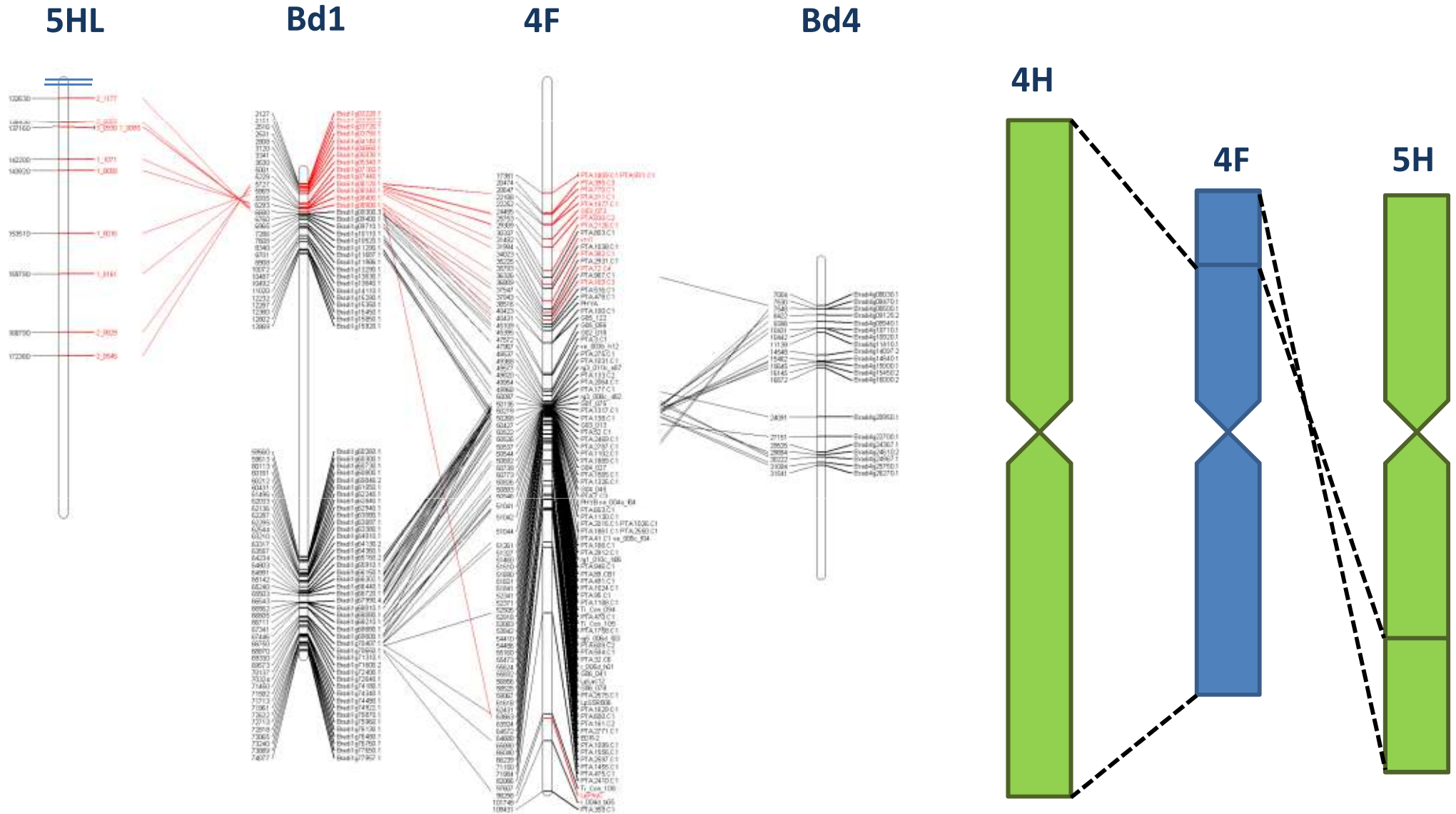
festuca_4F_vs_barley



*chromosome
rearrangements*



Ancestral translocation 4F/4H-5HL



Is it possible to introgress any segments from any part of genome?

Development of introgression lines

Fp ($2n=2x=14$) x Lm ($2n=4x=28$)



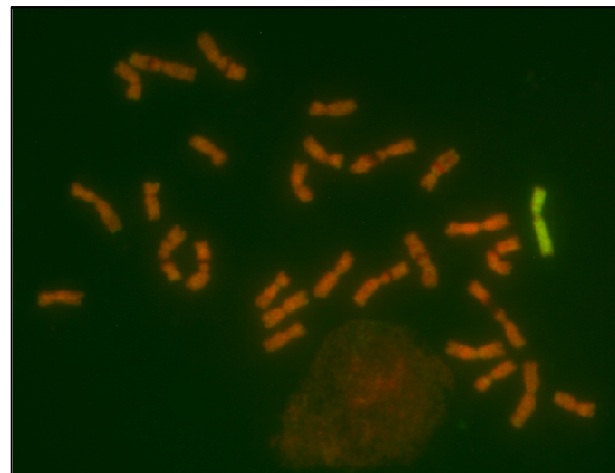
F1: ($2n=3x=21$, FpLmLm) x Lm ($2n=4x=28$)



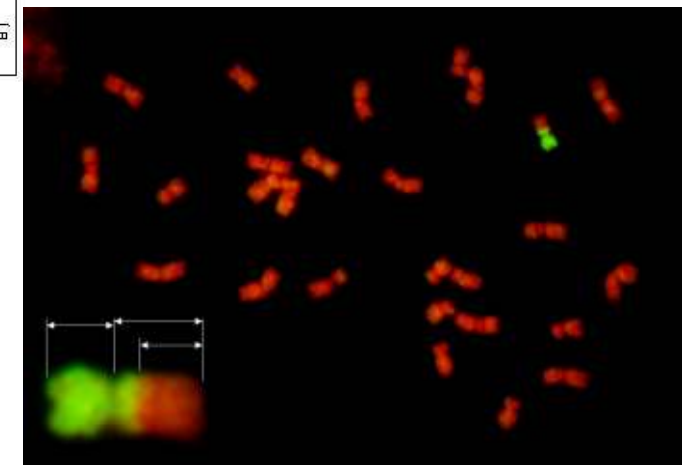
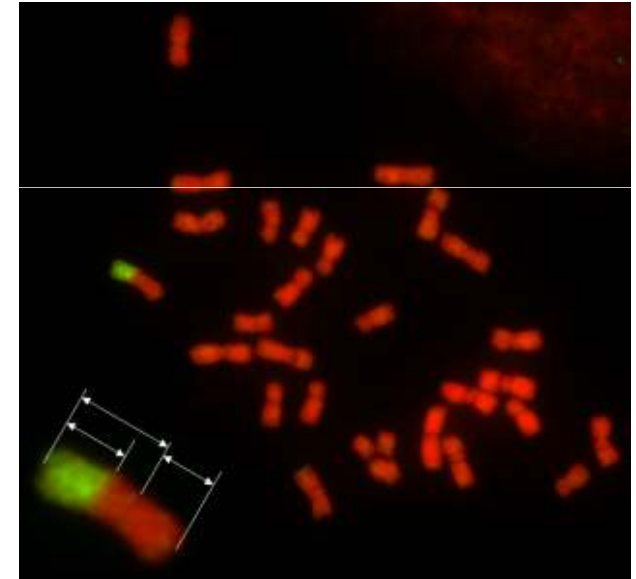
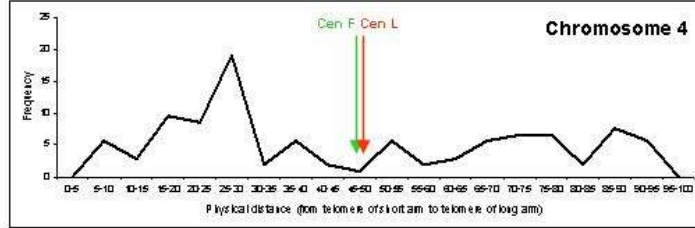
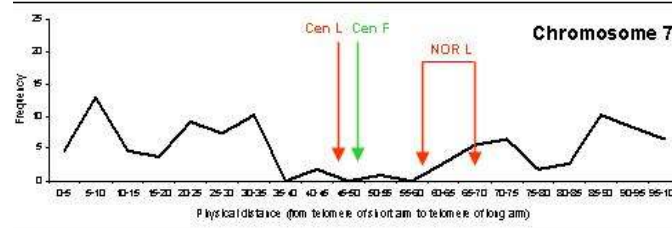
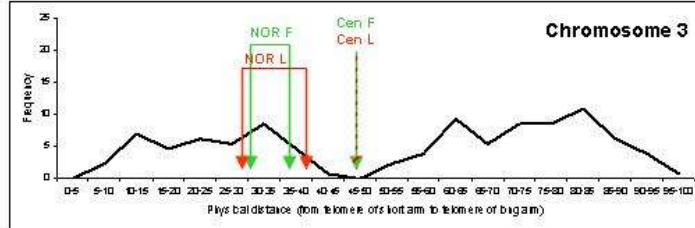
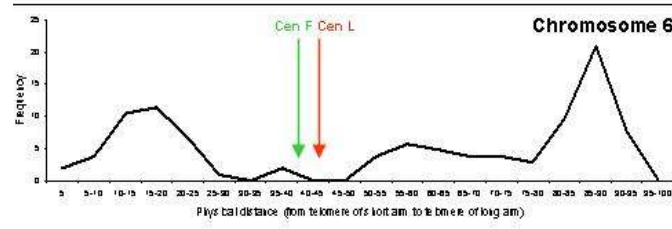
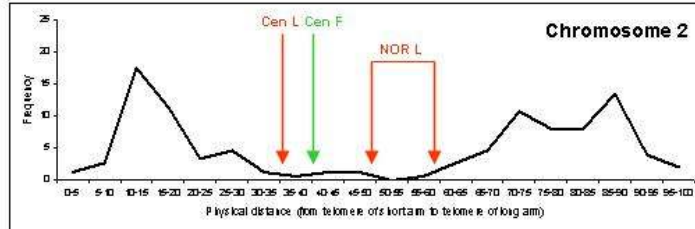
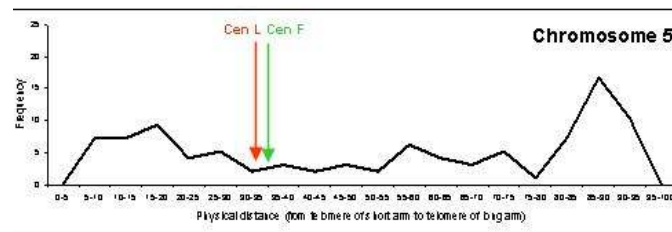
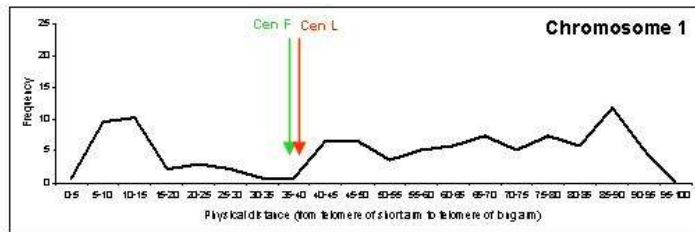
F2: ($2n=4x=28$, FpLmLmLm) x Lm ($2n=4x=28$)



F3 etc.



Distribution of homeologous recombination in *Festuca/Lolium*



Fundings

- 2001-2005, Grant Agency of Czech Academy of Science (S5038104): Molecular cytogenetic and cytometric methods for breeding of grasses and clovers
- 2007-2009, Czech Science Foundation (grant award 521/07/P479): Cytogenetic mapping of genome of meadow fescue (*Festuca pratensis* Huds.)
- 2007-2010, The National Agency for Agriculture Research (grant award NAZV QH71267): The development and use of DArT array for xFestulolium breeding
- 2011-2014, Czech Science Foundation (grant award 501/11/0504): Genome interactions in interspecific hybrids xFestulolium

acknowledgements

