

Details of spinach BAC library as a resource

Chunda Feng and James Correll
Department of Plant Pathology
University of Arkansas

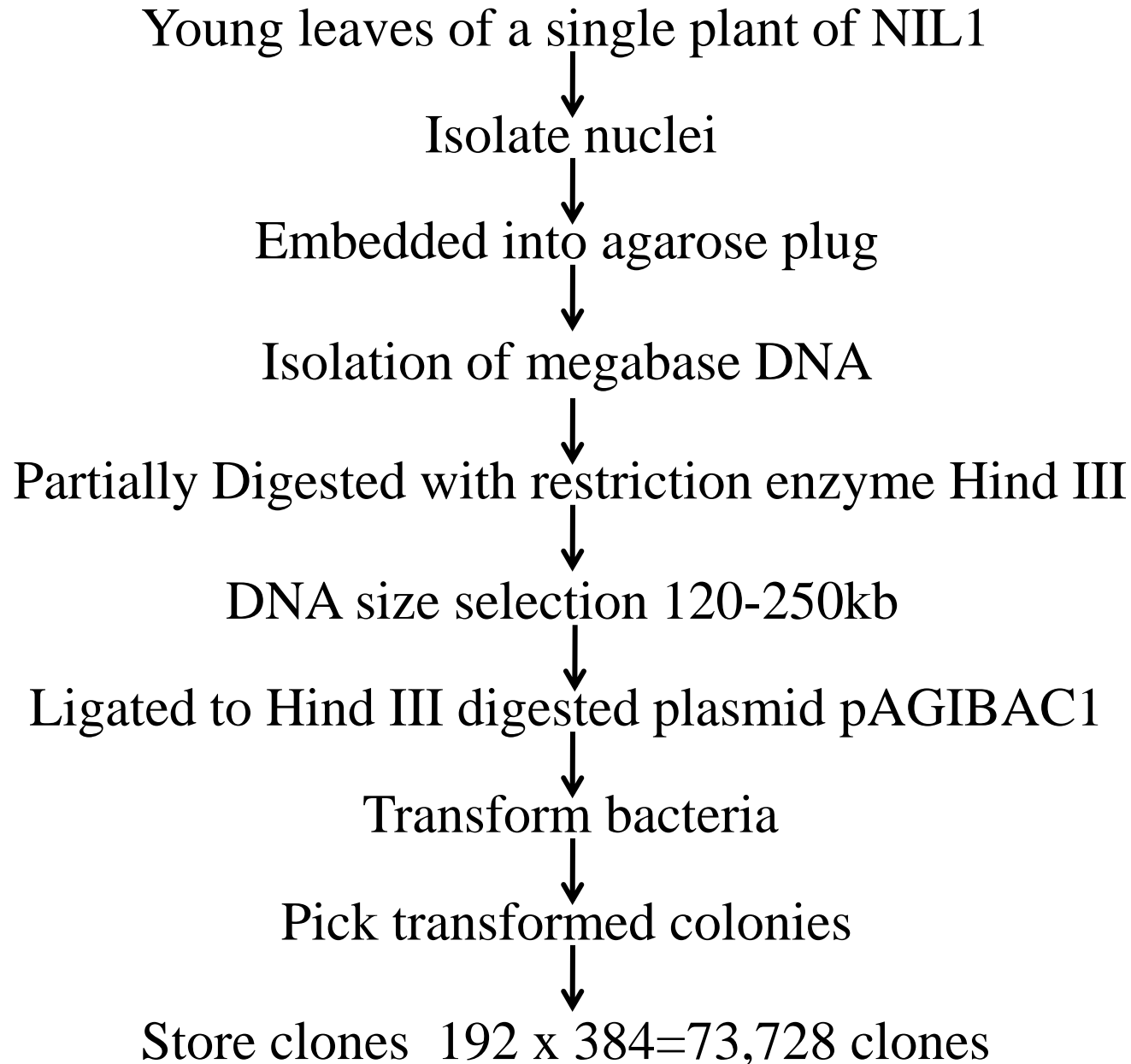
Spinach

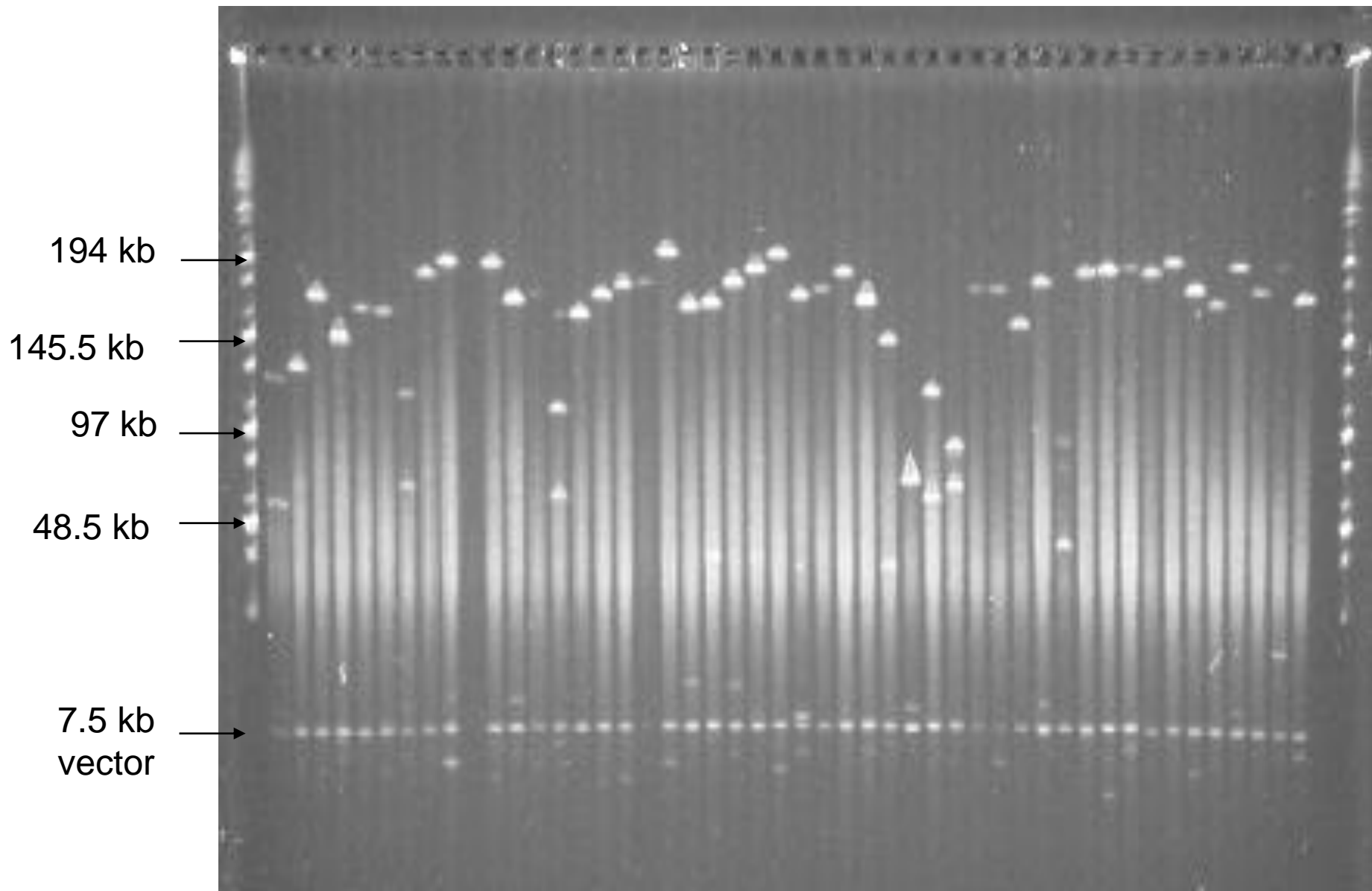
- Very nutritious
- Important crop
- Inheritances of a few simple traits
- Don't know the inheritances of many traits
- No genetic linkage map
- No sequence information available

BAC library

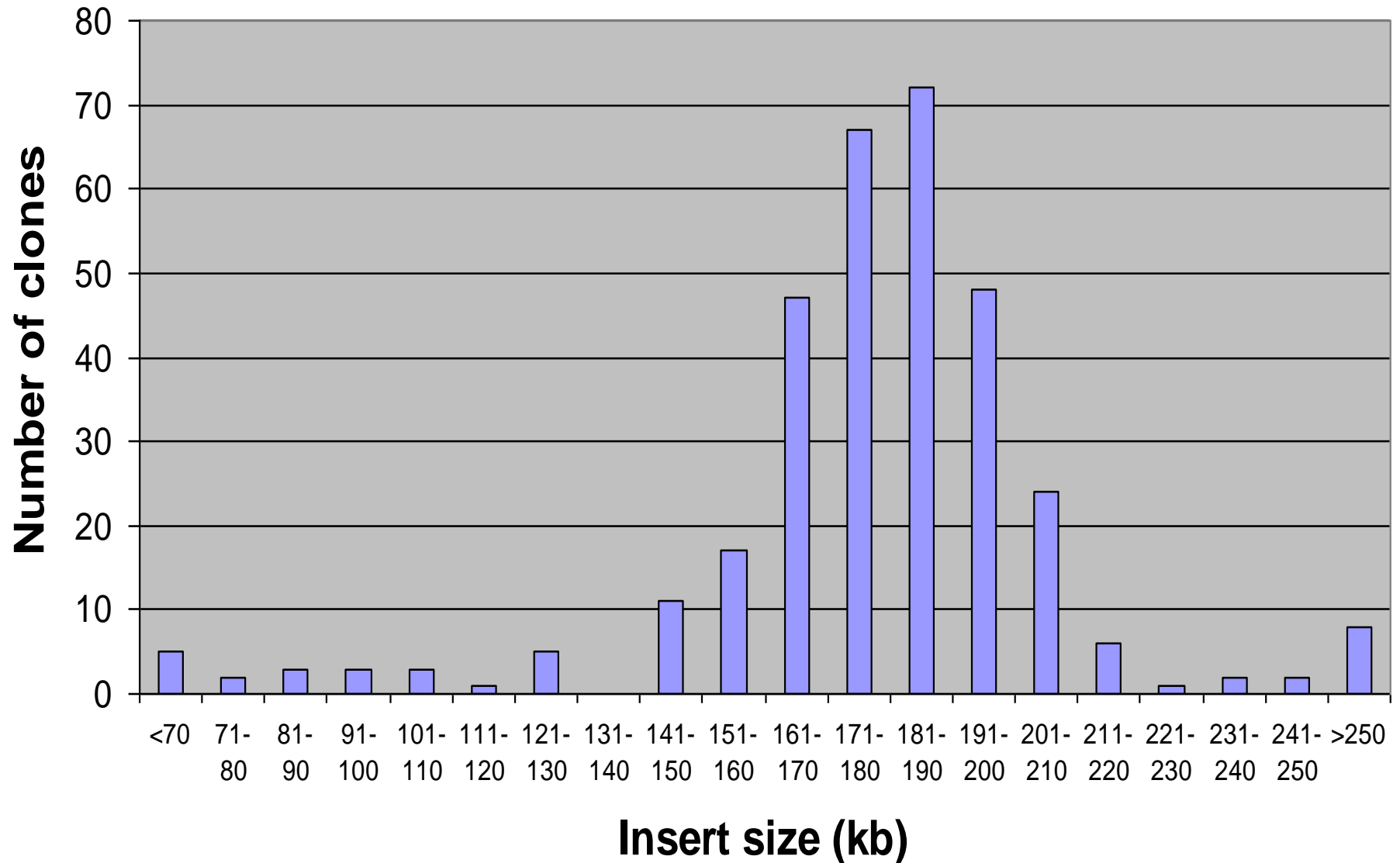
- Bacterial Artificial Chromosome library
- is a collection of bacterial clones that together represents the entire genome of a given organism

Spinach BAC library construction





***Spinacia oleracea* (Spinach)**
Average insert size 183 kb

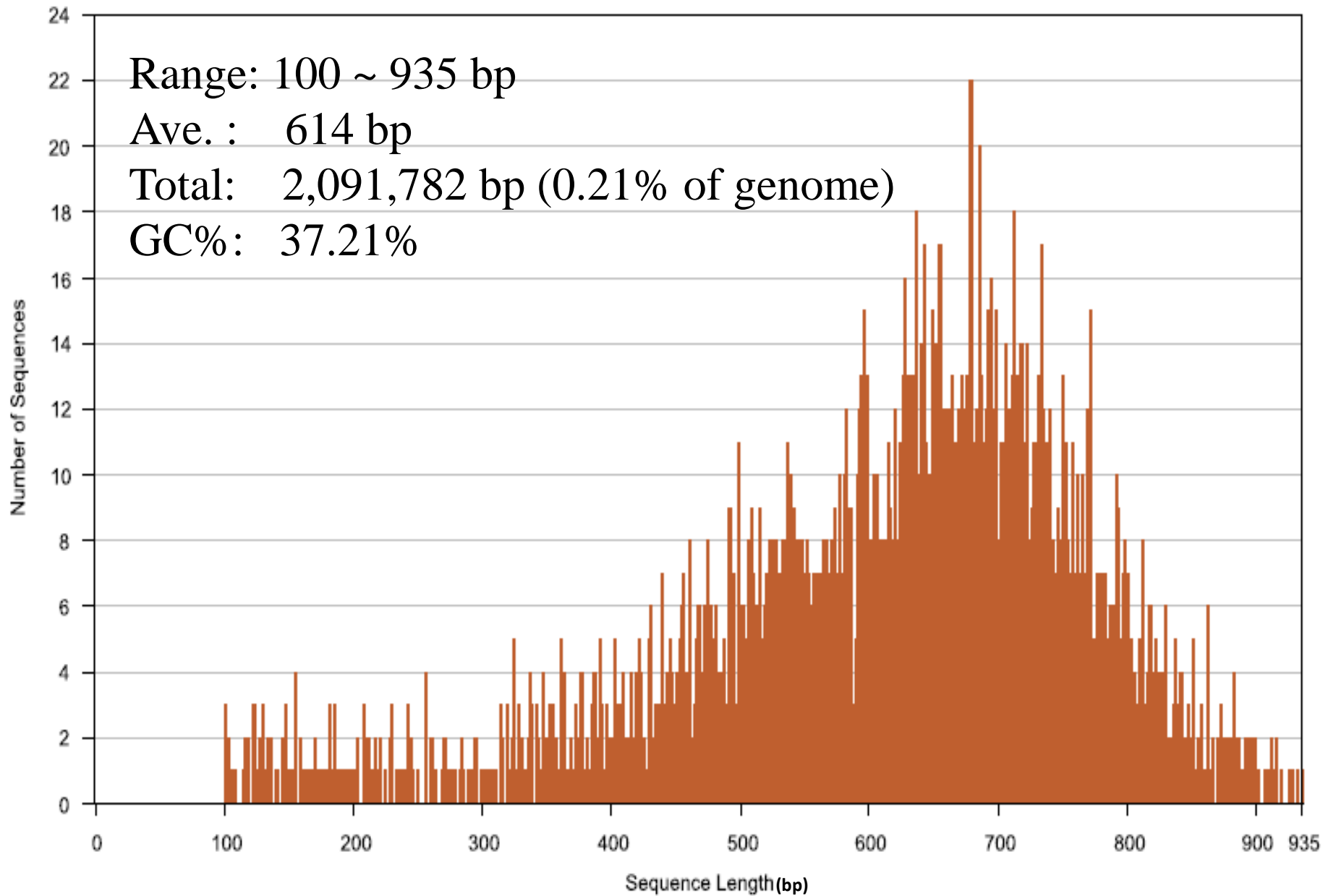


- Spinach genome size: 989 Mb
- BAC library: 73,728 clones
- Ave. insert size: 183 kb
- Coverage: ~13 x

$$\frac{73728 \times 183}{989 \times 1000} = 13.64$$

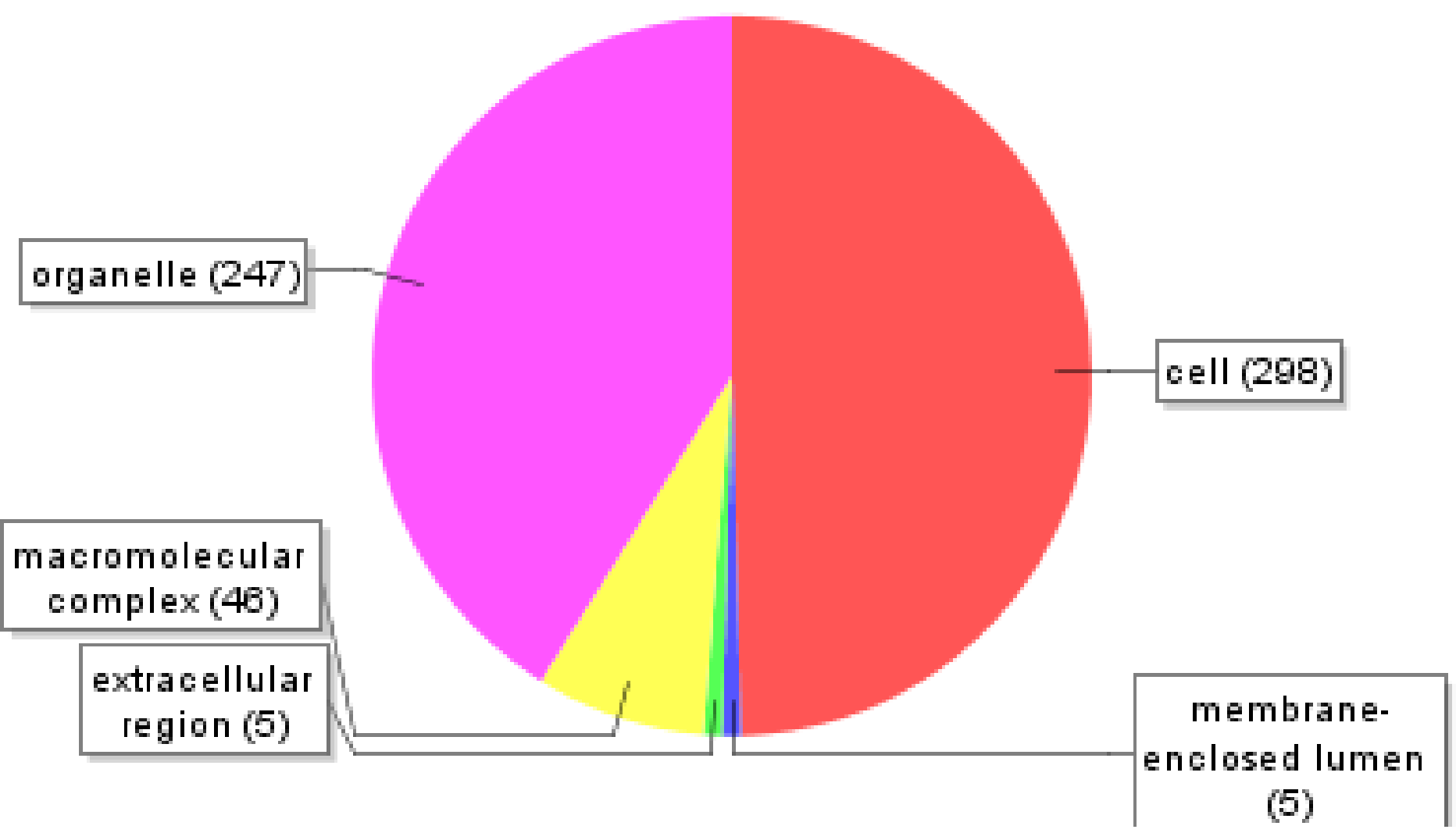
BAC end sequencing

- Five 384-well plates $5 \times 384 = 1920$ clones
- Sequenced with T7 and Sp6 primers
 $2 \times 1920 = 3840$ BAC ends
- Sequences
 - 101 clones - no sequences from both directions,
 - 134 clones – only forward sequences,
 - 99 clones – only reverse sequences
 - 3405 high quality BAC End Sequences (BESs)

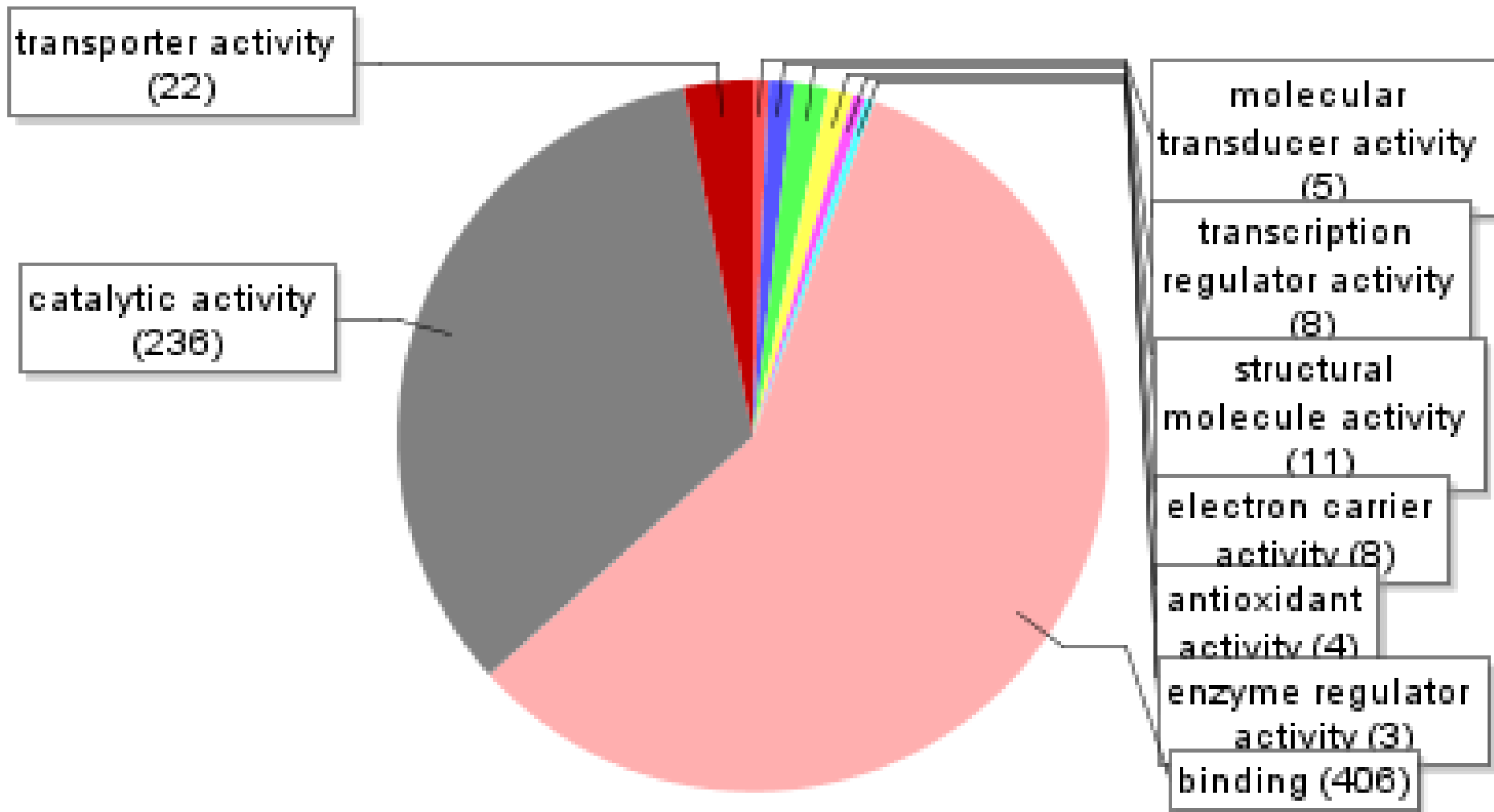


Repetitive DNA

| Type | Number of elements | Length | % of sequence |
|-----------------------------|--------------------|--------|---------------|
| Retroelements | 363 | 122159 | 5.84 |
| SINEs: | 0 | 0 | 0.00 |
| LINEs: | 30 | 5637 | 0.27 |
| RTE/Bov-B | 3 | 1550 | 0.07 |
| L1/CIN4 | 27 | 4087 | 0.20 |
| LTR elements: | 333 | 116522 | 5.57 |
| Ty1/Copia | 202 | 72181 | 3.45 |
| Gypsy/DIRS1 | 131 | 44341 | 2.12 |
| DNA transposons | 32 | 9364 | 0.45 |
| hobo-Activator | 11 | 1584 | 0.08 |
| Tc1-IS630-Pogo | 1 | 359 | 0.02 |
| En-Spm | 16 | 6292 | 0.30 |
| MuDR-IS905 | 3 | 604 | 0.03 |
| Tourist/Harbinger | 1 | 525 | 0.03 |
| Unclassified: | 2 | 353 | 0.02 |
| Total interspersed repeats: | | 131876 | 6.30 |



Cellular component



Molecular function

Homology-based spinach gene identification

General physiological related genes

DNA or RNA polymerases, helicases, endonucleases

ATPases, GTPases

dehydrogenases, synthases, kinases

transcription factors (zinc-finger protein, bZIP, AP2, MYB)

translational activator and initiation factors

phytases, pectinases, germination proteins

Reproduction related genes

sex determining protein MOG-3

sexual development activator VeA

ovule development protein

pollen-specific kinase partner protein

fertility restorer homologues

Environmental response related genes

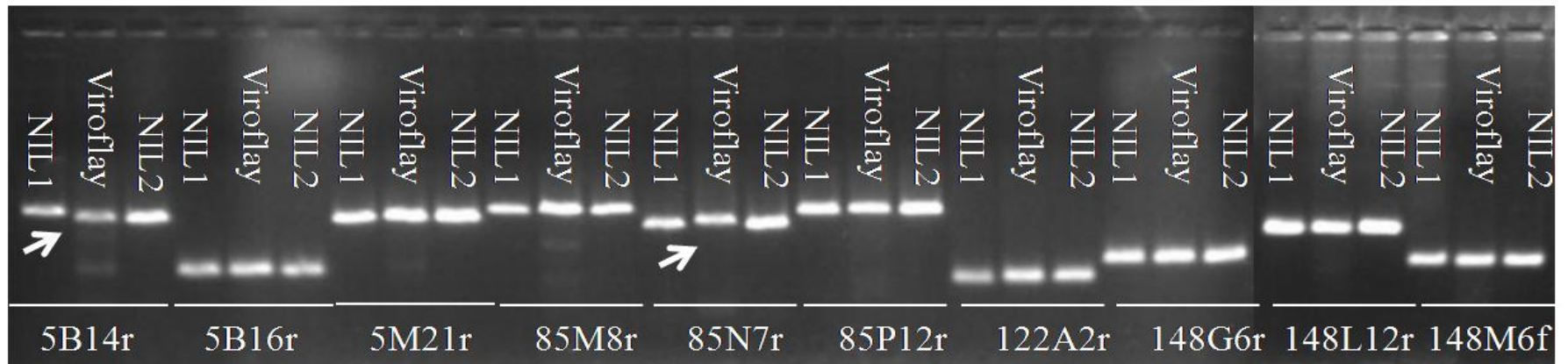
- 78 cytosolic heat shock proteins
- 11 cold acclimation proteins
- drought-induced proteins
- salt-induced proteins
- ethylene response factor

Disease resistance genes

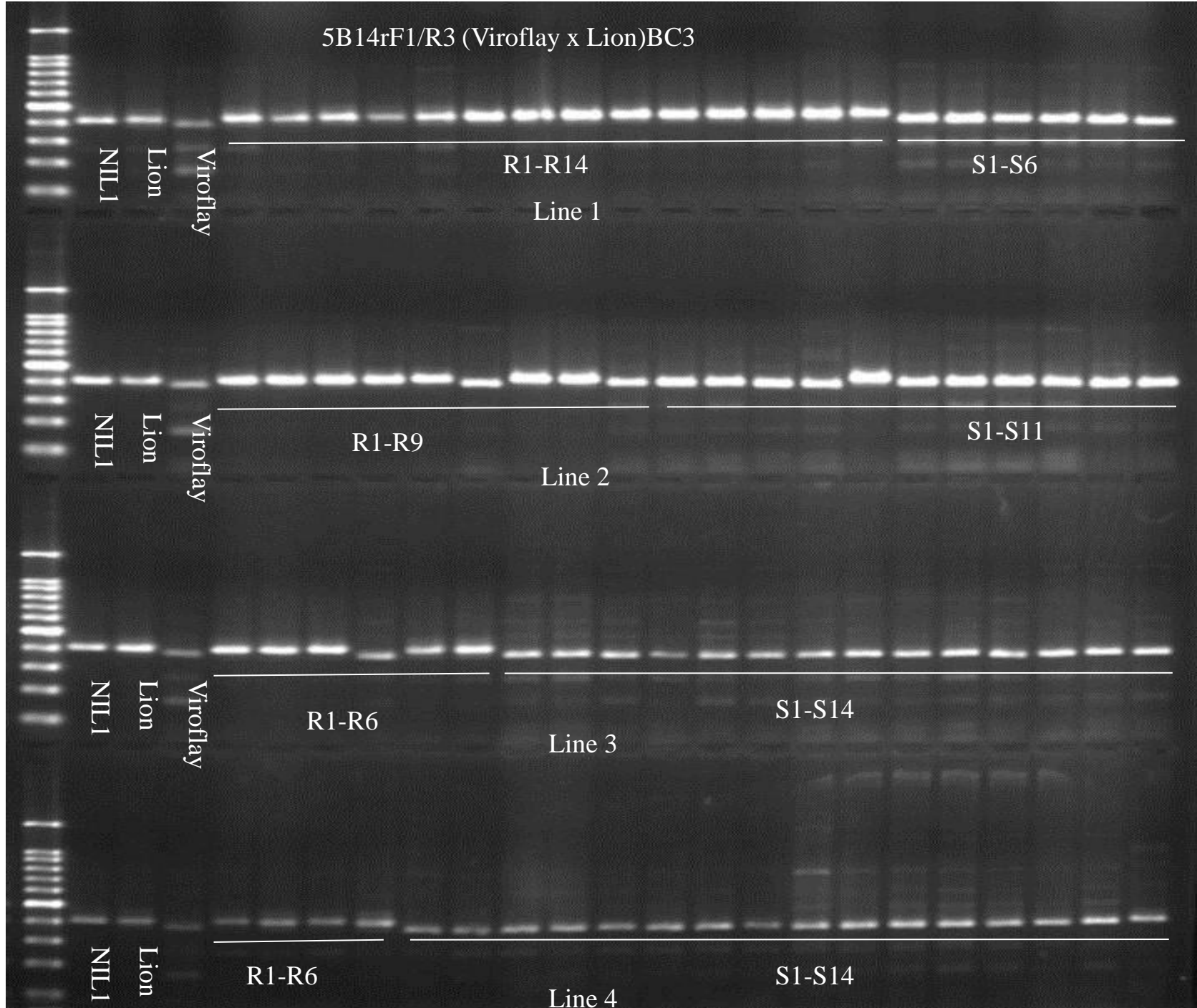
- 25 resistance gene analogs

| BES ID. | Putative function or domain | Genbank ID | E- value |
|----------------|-------------------------------------|-------------------|-----------------|
| 001K18.r.scf | LRR protein | ABH09321 | 3.79E-10 |
| 001K20.r.scf | Pathogenesis-induced protein | ACG76109 | 4.49E-03 |
| 001D11.r.scf | Avr9/Cf-9 rapidly elicited protein | AAG43545 | 3.31E-09 |
| 005A15.f.scf | Plant disease resistance protein | AAG13430 | 9.52E-24 |
| 005G15.f.scf | LRR receptor-like kinase | AAV33327 | 3.29E-03 |
| 005M21.r.scf | LRR family protein | ABF95661 | 4.09E-45 |
| 005J03.r.scf | LRR protein | ABH09321 | 3.06E-09 |
| 005B14.r.scf | NBS-LRR resistance protein | ABJ99599 | 7.30E-26 |
| 005B16.r.scf | LRR protein | ABH09321 | 2.61E-07 |
| 005N08.r.scf | avr9 elicitor response protein | BAD17751 | 1.14E-16 |
| 085I17.f.scf | Plant viral-response protein | NP_001147351 | 1.17E-03 |
| 085M08.r.scf | LRR protein | XP_002883938 | 4.47E-19 |
| 085N07.r.scf | NBS-LRR protein | ADB66335 | 1.44E-15 |
| 085P19.r.scf | Nucleotide binding | NP_850474 | 6.48E-16 |
| 085P12.r.scf | LRR protein | XP_002519373 | 3.93E-28 |
| 122O12.f.scf | Disease resistance protein | ACM89590 | 8.58E-06 |
| 122A02.r.scf | CC-NBS-LRR resistance protein | XP_002328558 | 1.51E-02 |
| 122J03.f.scf | Disease resistance protein | ACP30609 | 5.92E-20 |
| 148K01.r.scf | Barley stem rust resistance protein | AAX92856 | 2.55E-02 |
| 148M06.f.scf | Putative disease resistance protein | AAM12333 | 1.21E-10 |
| 148G16.r.scf | LRR protein | ACC91251 | 5.35E-19 |
| 148L12.f.scf | Plant disease resistance protein | ACP30609 | 4.86E-12 |
| 148P24.r.scf | NBS-LRR resistance protein | XP_002328559 | 8.63E-04 |

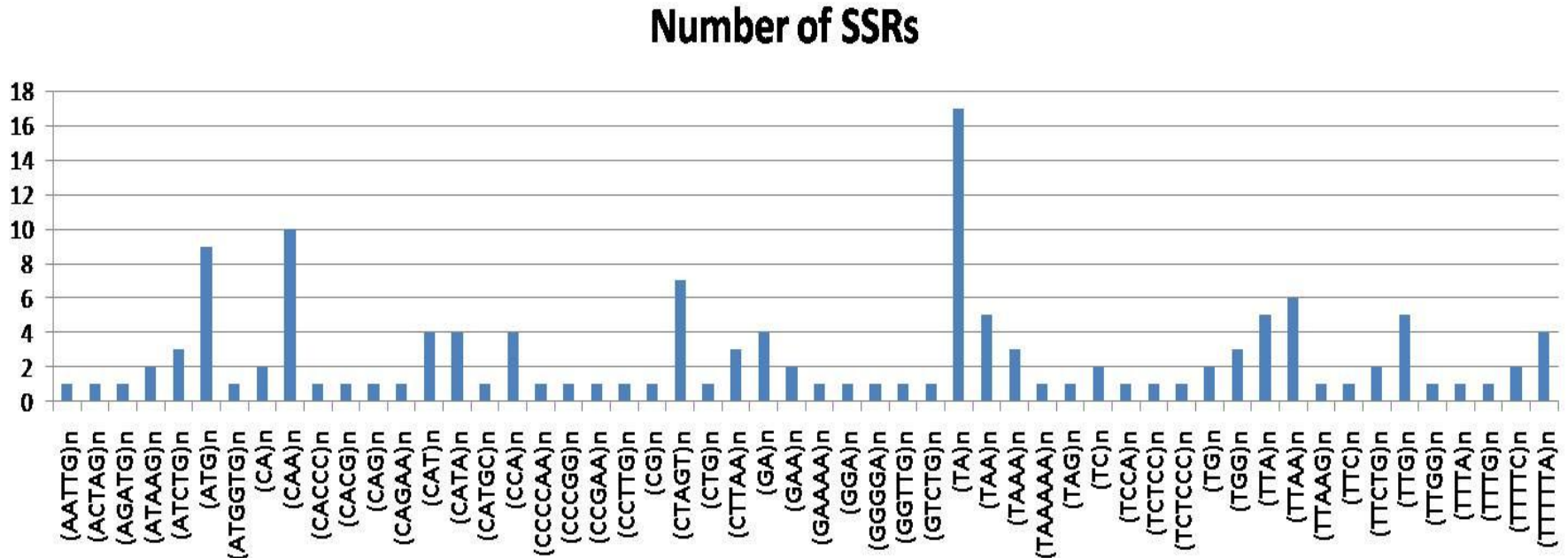
Testing 10 spinach RGA primers on NIL1, NIL2 and Viroflay



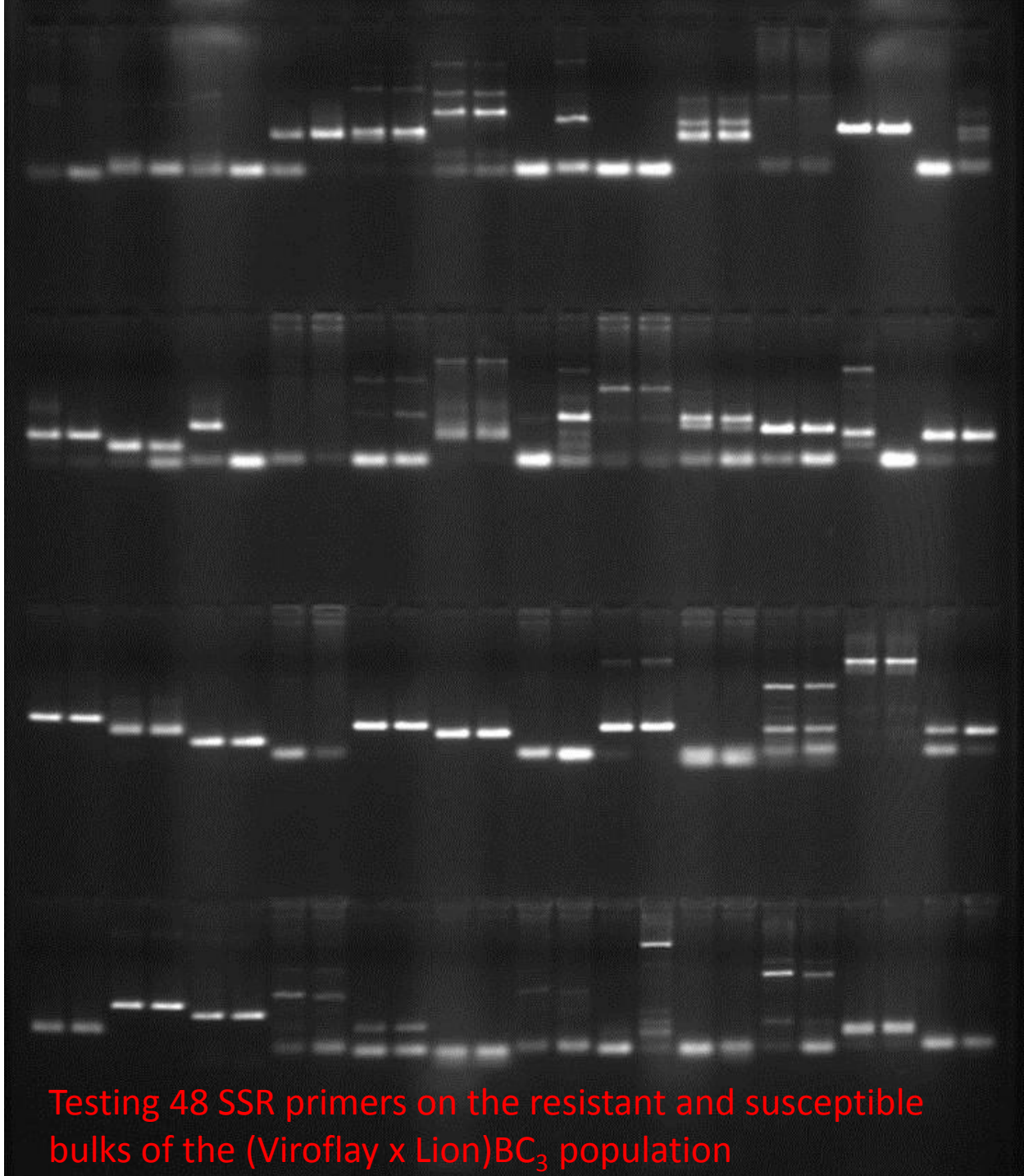
5B14rF1/R3 (Viroflay x Lion)BC3



SSR identification and marker development

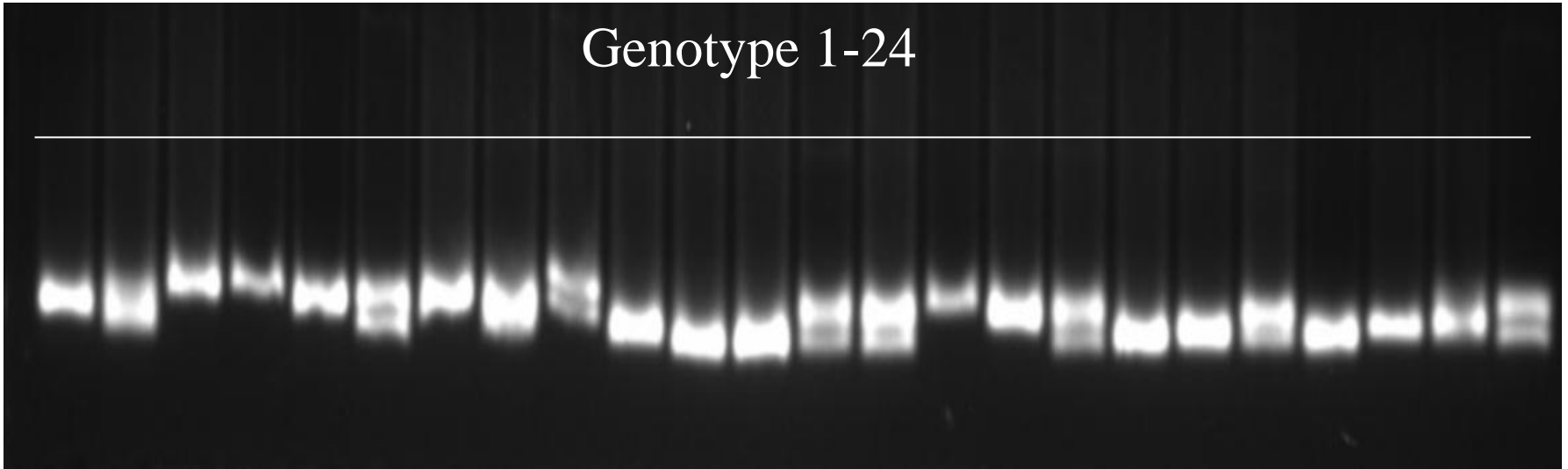


- 154 microsatellite or simple sequence repeats (SSRs) motifs
- di-, tri-, tetra-, penta-, and hexa- nucleotide repeats
- 100 pairs of SSR primers were developed

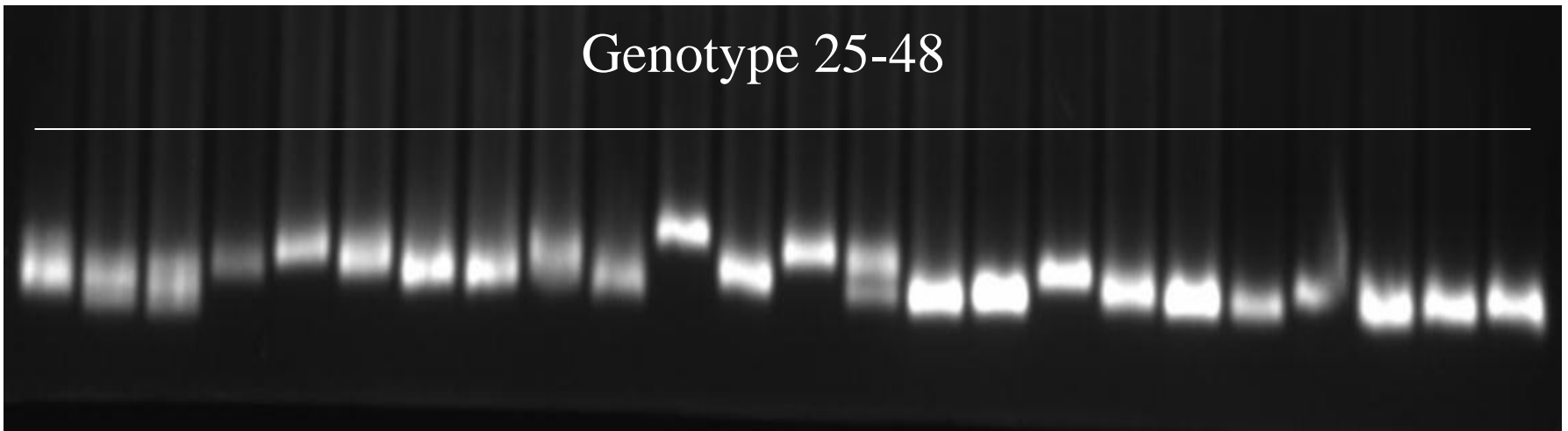


Diversity of spinach germplasm revealed by Sp026

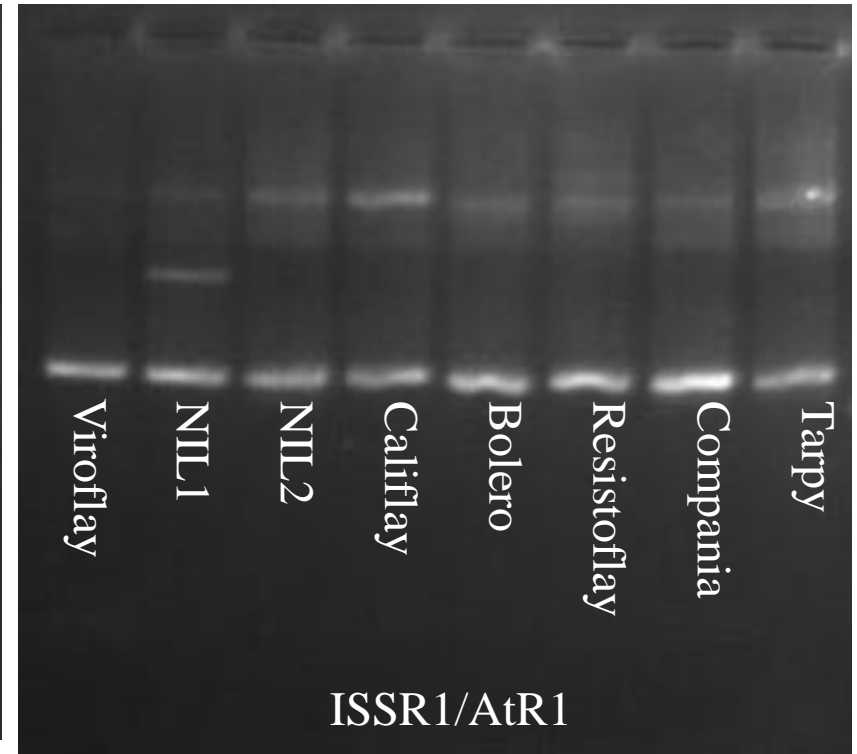
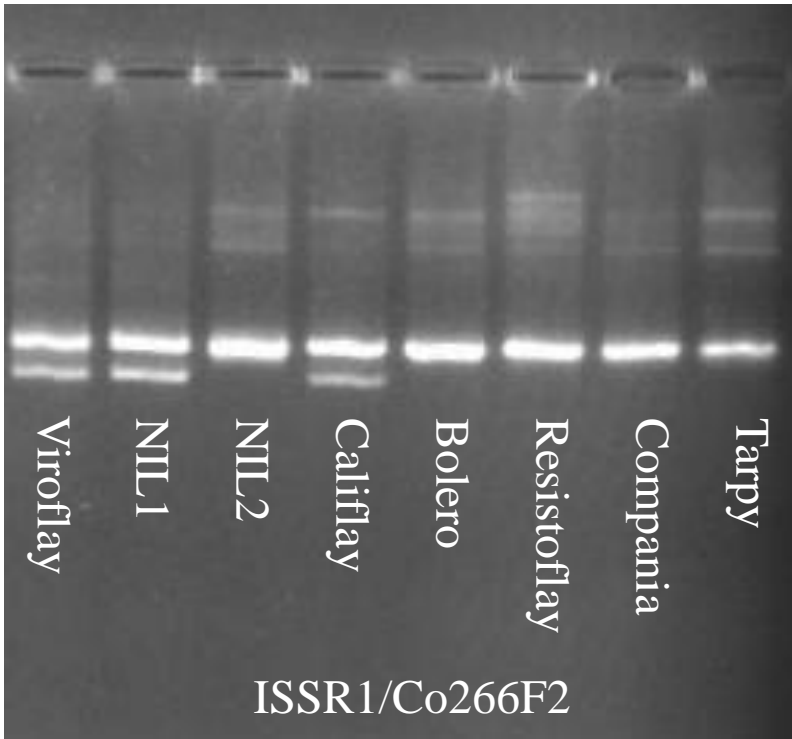
Genotype 1-24



Genotype 25-48



Retrotransposon-Microsatellite Amplified Polymorphism (REMAP)



Summary

- Constructed a high spinach BAC library
 - 73,728 clones
 - Ave. insert size 183kb
 - ~13 x coverage
- BAC end sequences
 - 20% BESs containing repetitive sequences
 - ~6% total length are repetitive sequences
 - Gene identification and annotation
 - Marker development (SSRs, IRAP, and REMAP)