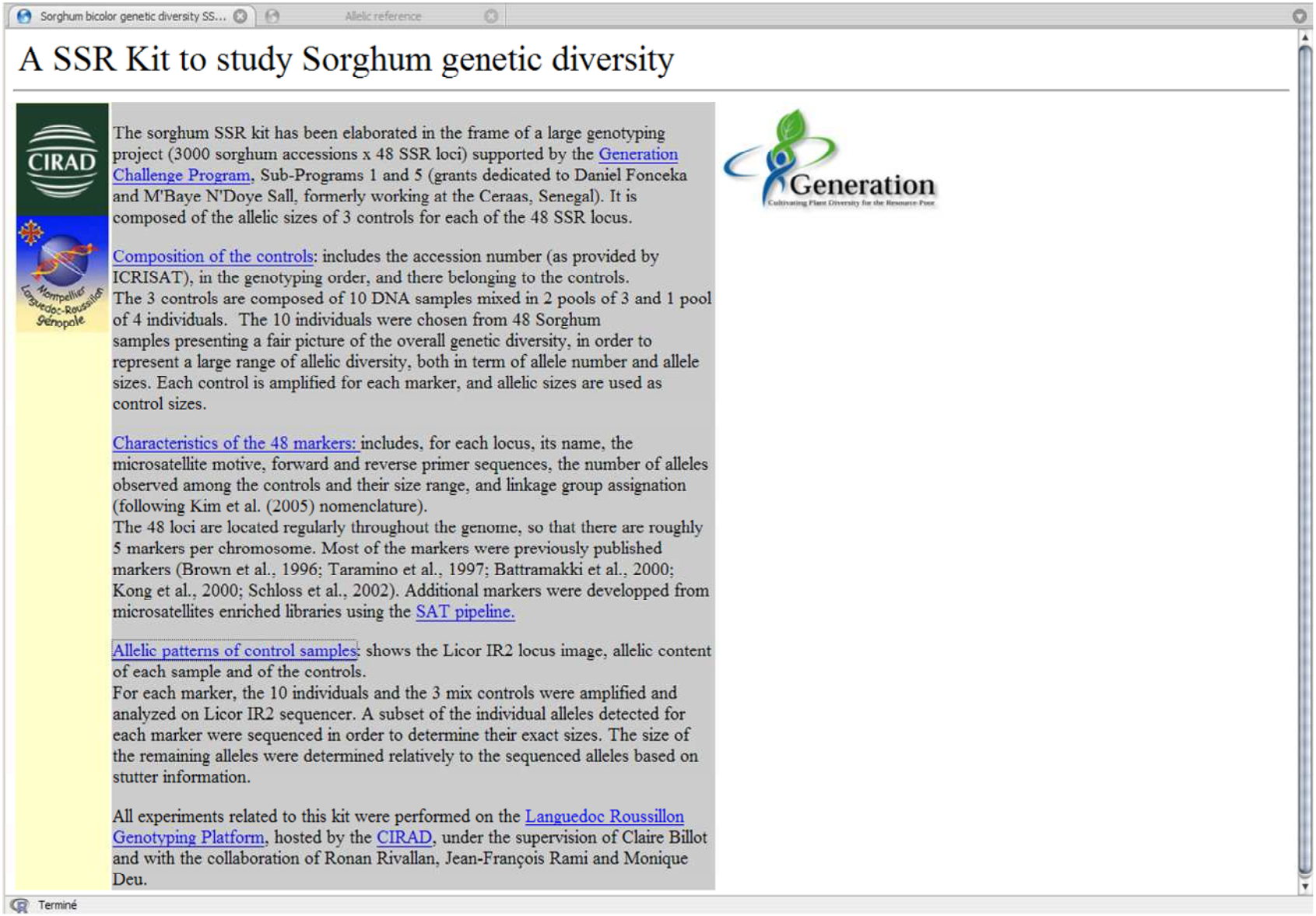
Sorghum SSR reference kit

- http://sat.cirad.fr/sat/sorghum_SSR_kit/


Sorghum bicolor genetic diversity SS...
Allelic reference
A SSR Kit to study Sorghum genetic diversity
CIRAD
Montpellier Languedoc-Roussillon Génopole
The sorghum SSR kit has been elaborated in the frame of a large genotyping project (3000 sorghum accessions x 48 SSR loci) supported by the Generation Challenge Program, Sub-Programs 1 and 5 (grants dedicated to Daniel Fonceka and M'Baye N'Doye Sall, formerly working at the Ceraas, Senegal). It is composed of the allelic sizes of 3 controls for each of the 48 SSR locus.
Composition of the controls: includes the accession number (as provided by ICRISAT), in the genotyping order, and there belonging to the controls.
The 3 controls are composed of 10 DNA samples mixed in 2 pools of 3 and 1 pool of 4 individuals. The 10 individuals were chosen from 48 Sorghum samples presenting a fair picture of the overall genetic diversity, in order to represent a large range of allelic diversity, both in term of allele number and allele sizes. Each control is amplified for each marker, and allelic sizes are used as control sizes.
Characteristics of the 48 markers: includes, for each locus, its name, the microsatellite motive, forward and reverse primer sequences, the number of alleles observed among the controls and their size range, and linkage group assignation (following Kim et al. (2005) nomenclature).
The 48 loci are located regularly throughout the genome, so that there are roughly 5 markers per chromosome. Most of the markers were previously published markers (Brown et al., 1996; Taramino et al., 1997; Battramakki et al., 2000; Kong et al., 2000; Schloss et al., 2002). Additional markers were developped from microsatellites enriched libraries using the SAT pipeline.
Allelic patterns of control samples: shows the Licor IR2 locus image, allelic content of each sample and of the controls.
For each marker, the 10 individuals and the 3 mix controls were amplified and analyzed on Licor IR2 sequencer. A subset of the individual alleles detected for each marker were sequenced in order to determine their exact sizes. The size of the remaining alleles were determined relatively to the sequenced alleles based on stutter information.
All experiments related to this kit were performed on the Languedoc Roussillon Genotyping Platform, hosted by the CIRAD, under the supervision of Claire Billot and with the collaboration of Ronan Rivallan, Jean-François Rami and Monique Deu.
Generation
Cultivating Plant Diversity for the Resource-Poor
Terminé

Sorghum bicolor genetic diversity SS... | Allelic reference | Sorghum SSR marker characteristics

Sorghum SSR marker characteristics

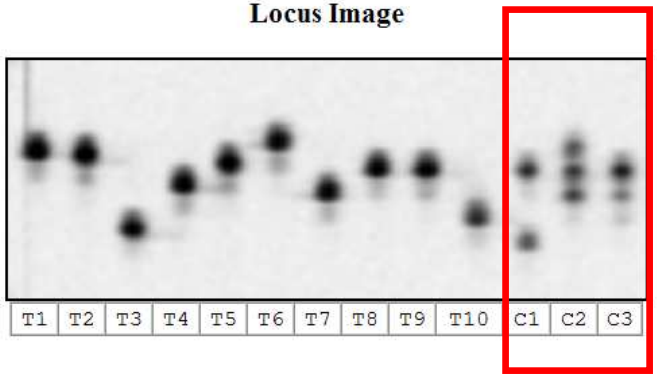
Marker_Name	Motive	Forward_Primer	Reverse_Primer	Reference	T°a	Size_Range	Allele_Nb	LG
gpsb067	(GT)10	TAGTCCATACACCTTTCA	TCTCTCACACACATTCTTC	unpublished, Agropolis-Cirad-Genoplante	49	170-180	6	8
gpsb069	(TC)12	CCCATAATACTTGACCTTC	ACTTACTCCCTCTGTCCC	unpublished, Agropolis-Cirad-Genoplante	50	189-197	4	9
gpsb089	(TG)9	ATCAGGTACAGCAGGTAGG	ATGCATCATGGCTGGT	unpublished, Agropolis-Cirad-Genoplante	50	165-173	4	1
gpsb123	(CA)7+(GA)5	ATAGATGTTGACGAAGCA	GTGGTATGGGACTGGA	unpublished, Agropolis-Cirad-Genoplante	50	288-296	5	8
gpsb148	(TC)3+(CA)5	CAACCACAAACCAAGAG	ATAGAAATGGGGTGGAG	unpublished, Agropolis-Cirad-Genoplante	50	135-147	4	5
gpsb151	(CT)12	ATACCAAGTTTCCTTTACCT	GTTGGGGGAGAGTTTT	unpublished, Agropolis-Cirad-Genoplante	50	106-128	7	4
ISEP0107	(TGG)4	GCCGTAACAGAGAAGGATGG	TTTCCGCTACCTCAAAAACC	unpublished, Icrisat	59	199-205	3	3
ISEP0310	(CCAAT)4	TGCCTTGTGCCTTGTTTATCT	GGATCGATGCCTATCTCGTC	unpublished, Icrisat	60	164-219	3	2
mSbCIR223	(AC)6	CGTTCCAATGACTTTTCTTC	GCCAATGTGGTGTGATAAAT	unpublished, Agropolis-Cirad-Genoplante	55	108-118	5	2
mSbCIR238	(AC)26	AGAAGAAAAGGGGTAAGAGC	CGAGAAACAATTACATGAACC	unpublished, Agropolis-Cirad-Genoplante	55	79-107	8	2
mSbCIR240	(TG)9	GTTCTTGGCCCTACTGAAT	TCACCTGTAACCCTGTCTTC	unpublished, Agropolis-Cirad-Genoplante	55	108-112	2	8
mSbCIR246	(CA)7.5	TTTTGTTGCACTTTTGAGC	GATGATAGCGACCACAAATC	unpublished, Agropolis-Cirad-Genoplante	55	98-100	2	5
mSbCIR248	(GT)7.5	GTTGGTCAGTGGTGGATAAA	ACTCCCATGTGCTGAATCT	unpublished, Agropolis-Cirad-Genoplante	56	89-101	4	10
mSbCIR262	(CATG)3.25	GCACCAAAATCAGCGTCT	CCATTTACCCGTGGATTAGT	unpublished, Agropolis-Cirad-Genoplante	57	208-220	4	7
mSbCIR276	(AC)9	CCCCAATCTAACTATTTGGT	GAGGCTGAGATGCTCTGT	unpublished, Agropolis-Cirad-Genoplante	53	230-234	4	3
mSbCIR283	(CT)8 (GT)8.5	TCCCTTCTGAGCTTGTAAAT	CAAGTCACTACCAAATGCAC	unpublished, Agropolis-Cirad-Genoplante	54	113-139	8	7
mSbCIR286	(AC)9	GCTTCTATACTCCCCTCCAC	TTTATGGTAGGATGCTCTGC	unpublished, Agropolis-Cirad-Genoplante	55	112-134	4	1
mSbCIR300	(GT)9	TTGAGAGCGGCGAGGTAA	AAAAGCCCAAGTCTCAGTGCTA	unpublished, Agropolis-Cirad-Genoplante	61	104-110	4	5
mSbCIR306	(GT)7	ATACTCTCGTACTCGGCTCA	ATACTCTCGTACTCGGCTCA	unpublished, Agropolis-Cirad-Genoplante	56	120-124	3	1
mSbCIR329	(AC)8.5	GCAGAACATCACTCAAAGAA	GCAGAACATCACTCAAAGAA	unpublished, Agropolis-Cirad-Genoplante	55	109-117	5	10
sb4-72	(AG)16	TGCCACCACTCTGGAAAAGGCTA	CTGAGGACTGCCCCAAATGTAGG	Brown et al., 1996	55	183-203	7	9
sb5-206	(AC)13/(AG)20	ATTCATCATCCTCATCCTCGTAGAA	AAAAACCAACCCGACCCACTC	Brown et al., 1996	55	106-146	10	6
sb6-84	(AG)14	CGCTCTCGGGATGAATGA	TAACGGACCACTAACAAATGATT	Brown et al., 1996	55	183-217	7	2
SbAG-B02	(AG)35	CTCTGATATGTCGTTGTGCT	ATAGAGAGGATAGCTTATAGCTCA	Taramino et al., 1997	55	96-154	6	5
Xcup02	(GCA)6	GACGCAGCTTTGCTCCTATC	GTCCAACCAACCCACGTATC	Schloss et al., 2002	54	192-204	4	6
Xcup11	(GCTA)4	TACCGCCATGTCATCATCAG	CGTATCGCAAGCTGTGTTTG	Schloss et al., 2002	54	165-172	2	3
Xcup14	(AG)10	TACATCACAGCAGGGACAGG	CTGGAAAGCCGAGCAGTATG	Schloss et al., 2002	54	211-225	4	3
Xcup53	(TTTA)5	GCAGGAGTATAGGCAGAGGC	CGACATGACAAGCTCAAACG	Schloss et al., 2002	54	186-198	3	1
Xcup61	(CAG)7	TTAGCATGTCCACCACAACC	AAAGCAACTCGTCTGATCCC	Schloss et al., 2002	54	198-201	2	3
Xcup62	(GAA)6	CGAGAAGATCGAGAGAACCC	TGAAGACGACGACGACAGAC	Schloss et al., 2002	54	190-193	2	1
Xcup63	(GGATGC)4	GTAAAGGGCAAGGCAACAAG	GCCCTACAAAATCTGCAAGC	Schloss et al., 2002	54	133-145	3	2
Xtxp010	(CT)14	ATACTATCAAGAGGGGAGC	AGTACTAGCCACACGTCAC	Kong et al., 2000	50	135-151	4	6
Xtxp012	(CT)22	AGATCTGGCGGCAACG	AGTCACCCATCGATCATC	Kong et al., 2000	55	181-205	9	4
Xtxp015	(TC)16	CACAAACACTAGTGCCTTATC	CATAGACACCTAGGCCATC	Kong et al., 2000	55	199-223	6	10
Xtxp021	(AG)18	GAGCTGCCATAGATTTGGTCG	ACCTCGTCCCACCTTTGTTG	Kong et al., 2000	60	169-199	7	4
xtxp040	(GGA)7	CAGCAACTTGCACTTGTC	GGGAGCAATTTGGCACTAG	Kong et al., 2000	55	129-141	4	5
Xtxp057	(GT)21	GGAACTTTTGACGGGTAGTGC	CGATCGTGATGTCCCAATC	Kong et al., 2000	55	223-257	8	9
Xtxp114	(AGG)8	CGTCTTCTACCGCGTCCT	CATAATCCCACTCAACAATCC	Kong et al., 2000	50	211-217	3	3
Xtxp136	(GCA)5	GCGAATAGCATCTTACAACA	ACTGATCATTGGCAGGAC	Kong et al., 2000	55	240-243	2	10
Xtxp141	(GA)23	TGTATGGCCTAGCTTATCT	CAACAAGCCAACCTAAA	Kong et al., 2000	55	135-167	9	7
Xtxp145	(AG)22	GTTCCTCCTGCCATTACT	CTTCCGCACATCCAC	Kong et al., 2000	55	208-244	8	9
Xtxp265	(GAA)19	GTCTACAGGCGTGCAAATAAAA	TTACCATGCTACCCCTAAAAGTGG	Kong et al., 2000	55	186-234	7	9
Xtxp273	(TTG)20	GTACCCATTTAAATTGTTTGCAGTAG	CAGAGGAGGAGGAAGAGAAGG	Kong et al., 2000	55	169-193	6	8
Xtxp278	(TTG)12	GGGTTTCAACTCTAGCCTACCGAACTTCCT	ATGCCTCATCATGGTTCGTTTTGCTT	Kong et al., 2000	50	243-252	3	5
Xtxp295	(TC)19	AAATCATGCATCCATGTTCGTCTTC	CTCCCGCTACAAGAGTACATTCATAGCTTA	Kong et al., 2000	55	153-183	8	5

Terminé

Sorghum bicolor genetic diversity SS... | Allelic reference

mSbCIR329

Locus Image



Allelic content in the samples

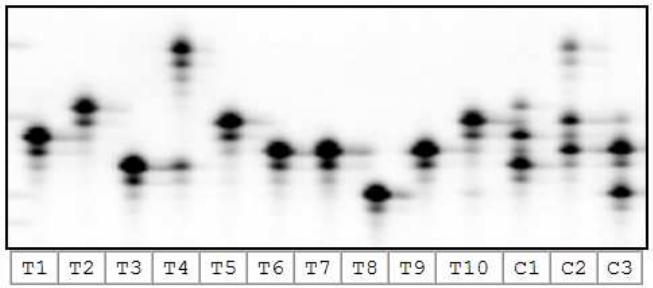
DNA_Sample	Control	Allele_Size
T1	C1	115
T2	C1	115
T3	C1	109
T4	C2	113
T5	C2	115
T6	C2	117
T7	C3	113
T8	C3	115
T9	C3	115
T10	C3	111

Allelic content in the controls

Control	Allele_Size
C1	109,115
C2	113,115,117
C3	111,113,115

sb4-72

Locus Image



Allelic content in the samples

DNA_Sample	Control	Allele_Size
T1	C1	191
T2	C1	195
T3	C1	187
T4	C2	187
T4	C2	203
T5	C2	193
T6	C2	189
T7	C3	189
T8	C3	183
T9	C3	189
T10	C3	193

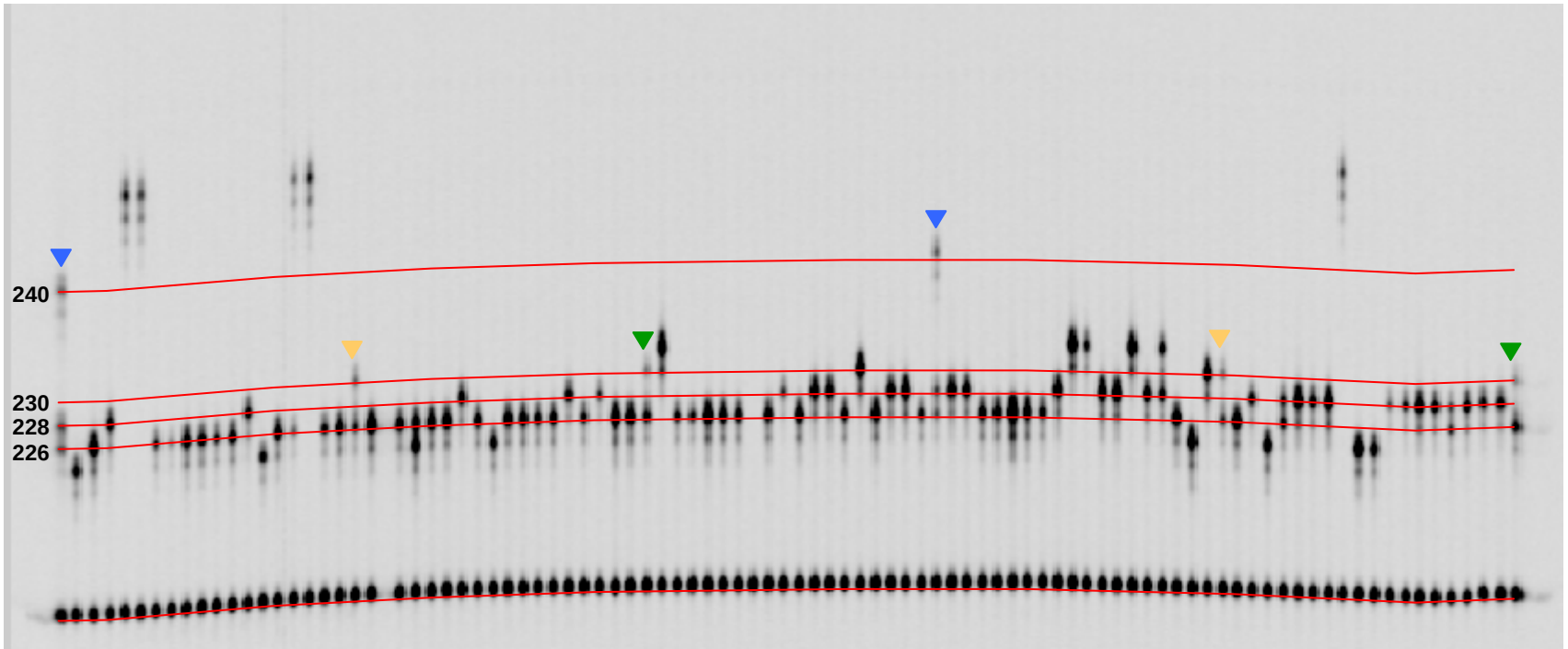
Allelic content in the controls

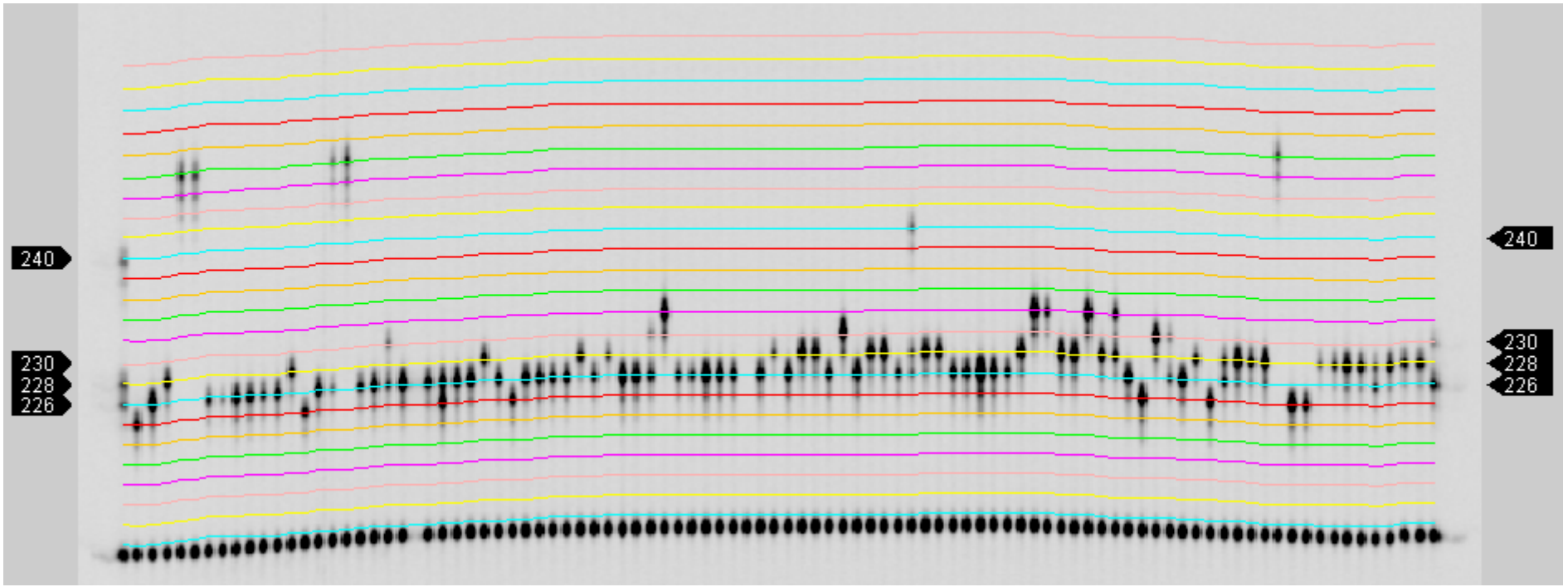
Control	Allele_Size
C1	187,191,195
C2	187,189,193,203
C3	183,189

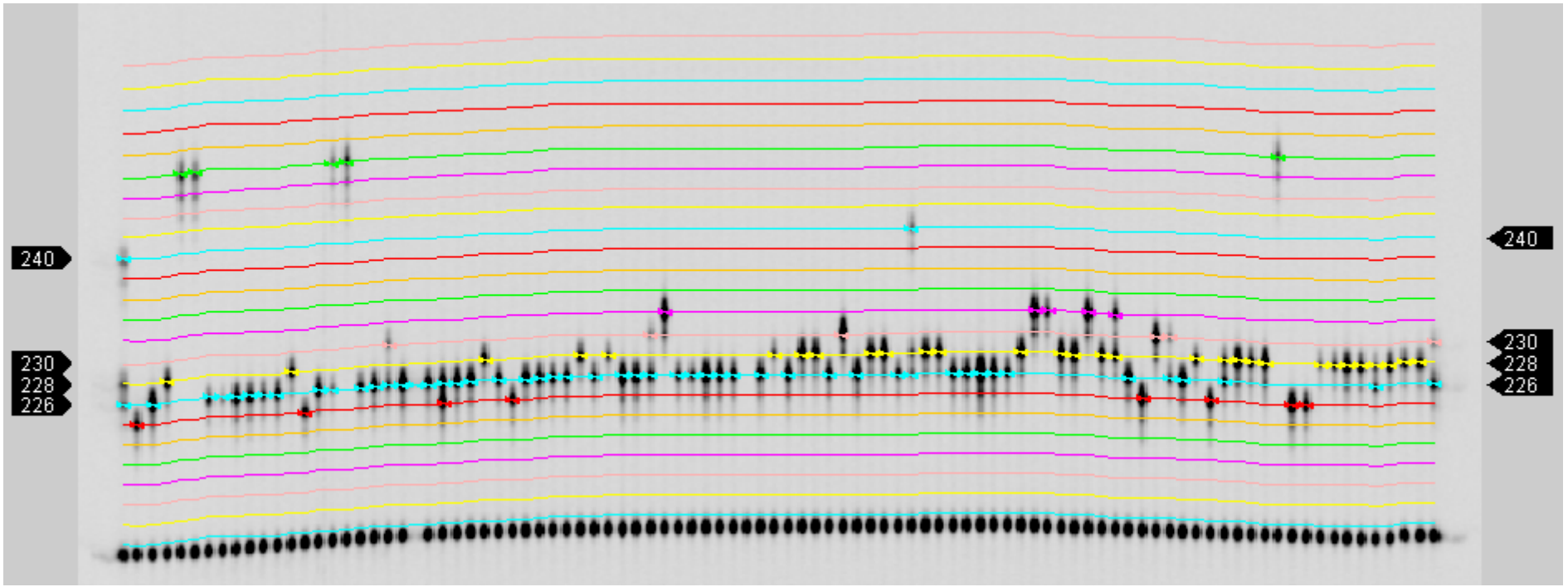
Terminé

Control		A	B	C
Xcup14	226	1	1	1
	228	1		
	230		1	1
	240	1		

Allelic control
Softwares
-SAGA software use
-AlleloBin




240
230
228
226
240
230
228
226


240
230
228
226
240
230
228
226

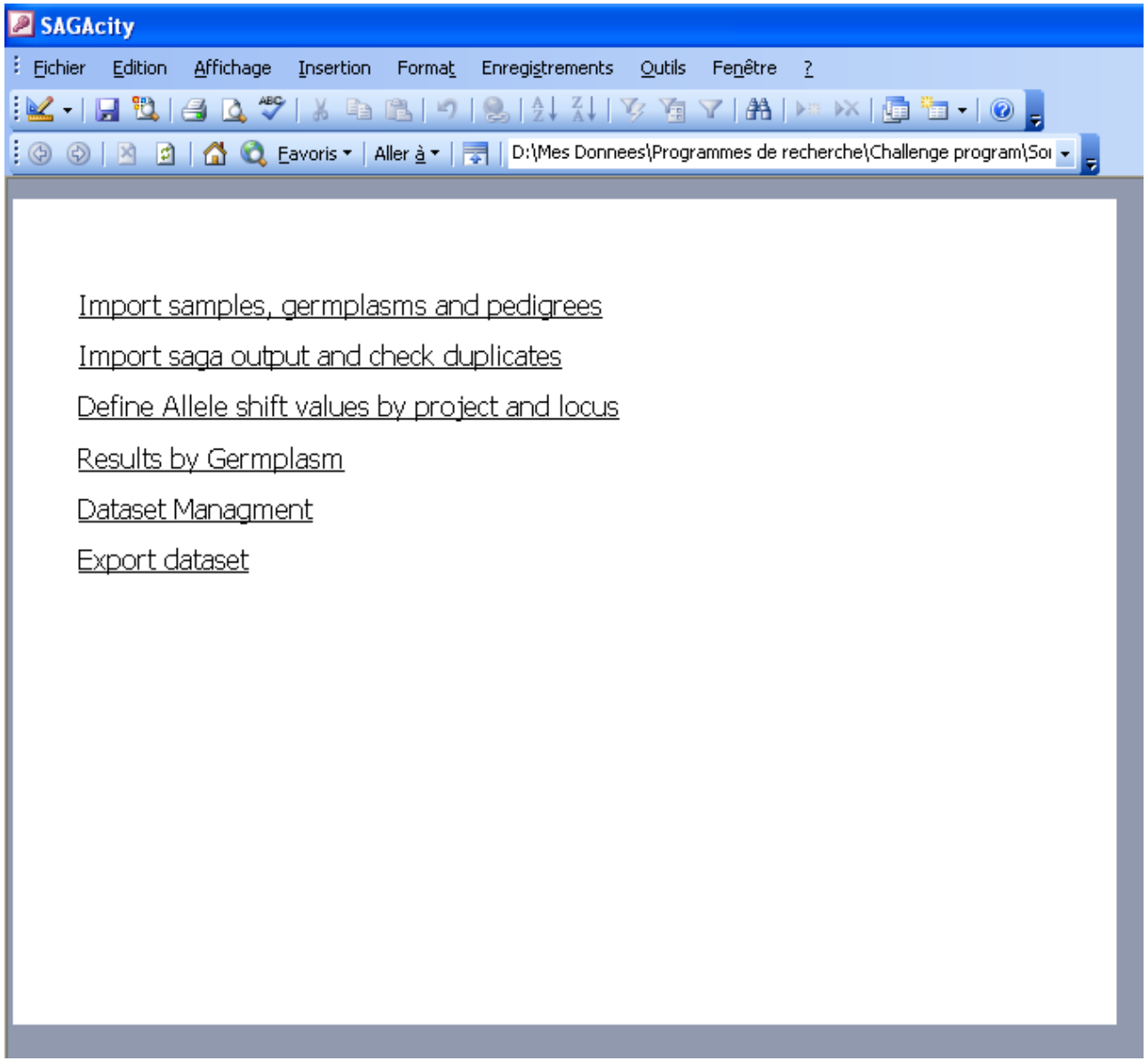
Sorghum Dataset – Merging Data

- At Cirad:
  - No LIMS available
  - Construction of an access db (SAGAcity) which enables the storage of germplam-sample correspondance, genotyping data, versionning and cleaning of the data
- At Icrisat:
  - LIMS
  - Extraction of a final dataset, including the control samples
- Between labs:
  - Performed in SagaCity, based on control bp sizes
- Submission to GCP of the dataset that is analysed
  - i.e. with less than missing data

SAGAcity

- Access db, oriented towards SAGA outputs, but adaptable (used to import data obtained from Icrisat)
- Build by JF Rami (Cirad) to extract specific datasets
  - Dealing with germplams-accession-DNA extraction
  - Enables the merging of datasets
  - Control the level of experiment repeats and user input on the quality of the data

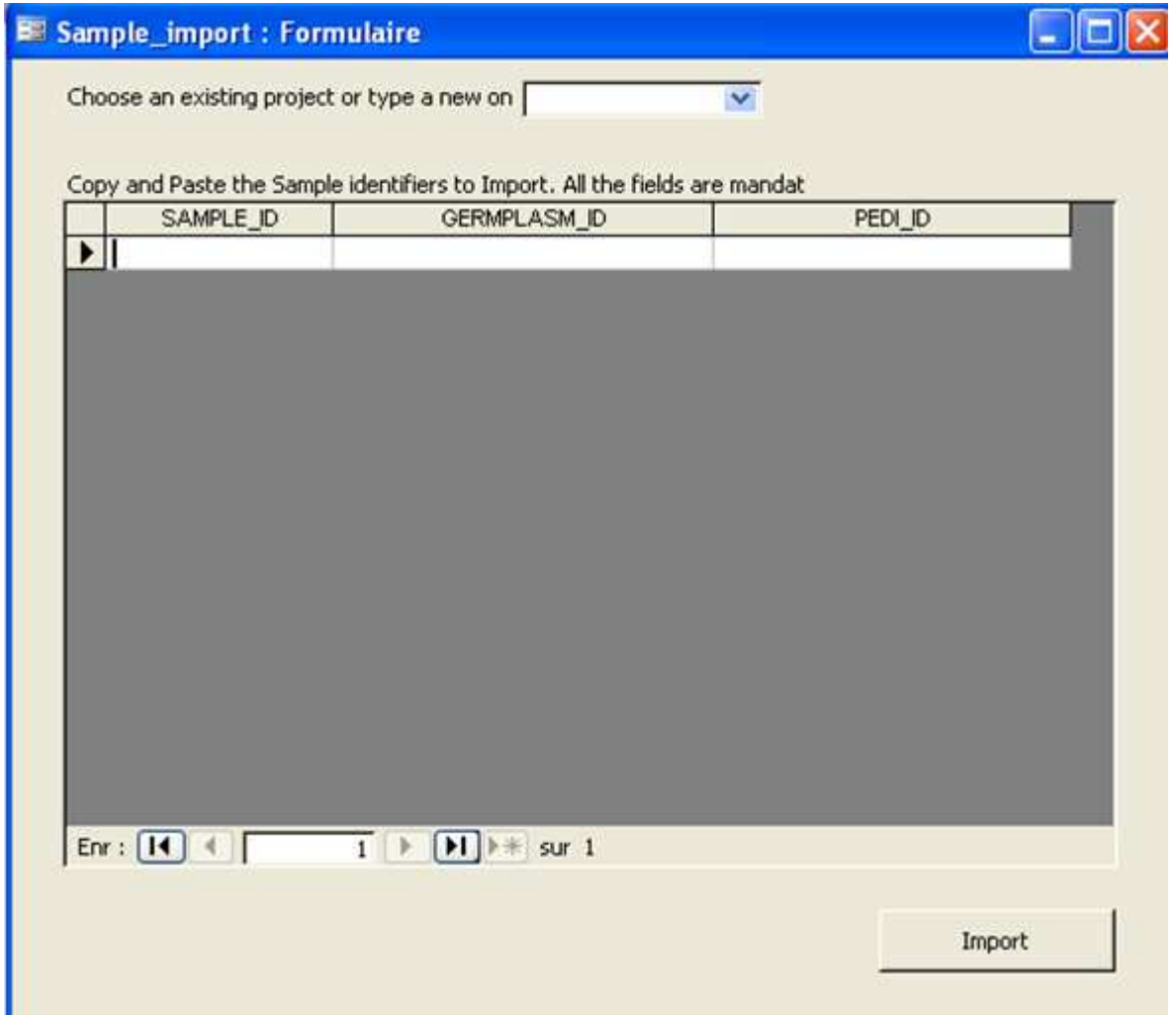
SagaCity – Main menu



SAGAcity – Defining the samples

Importing samples

- Sample_ID: DNA sample
- Germplasm_ID: seeds, etc…
  - The spirit of the db is to synthetise what has been obtained by germplasm
- Pedigree_ID = accession ID



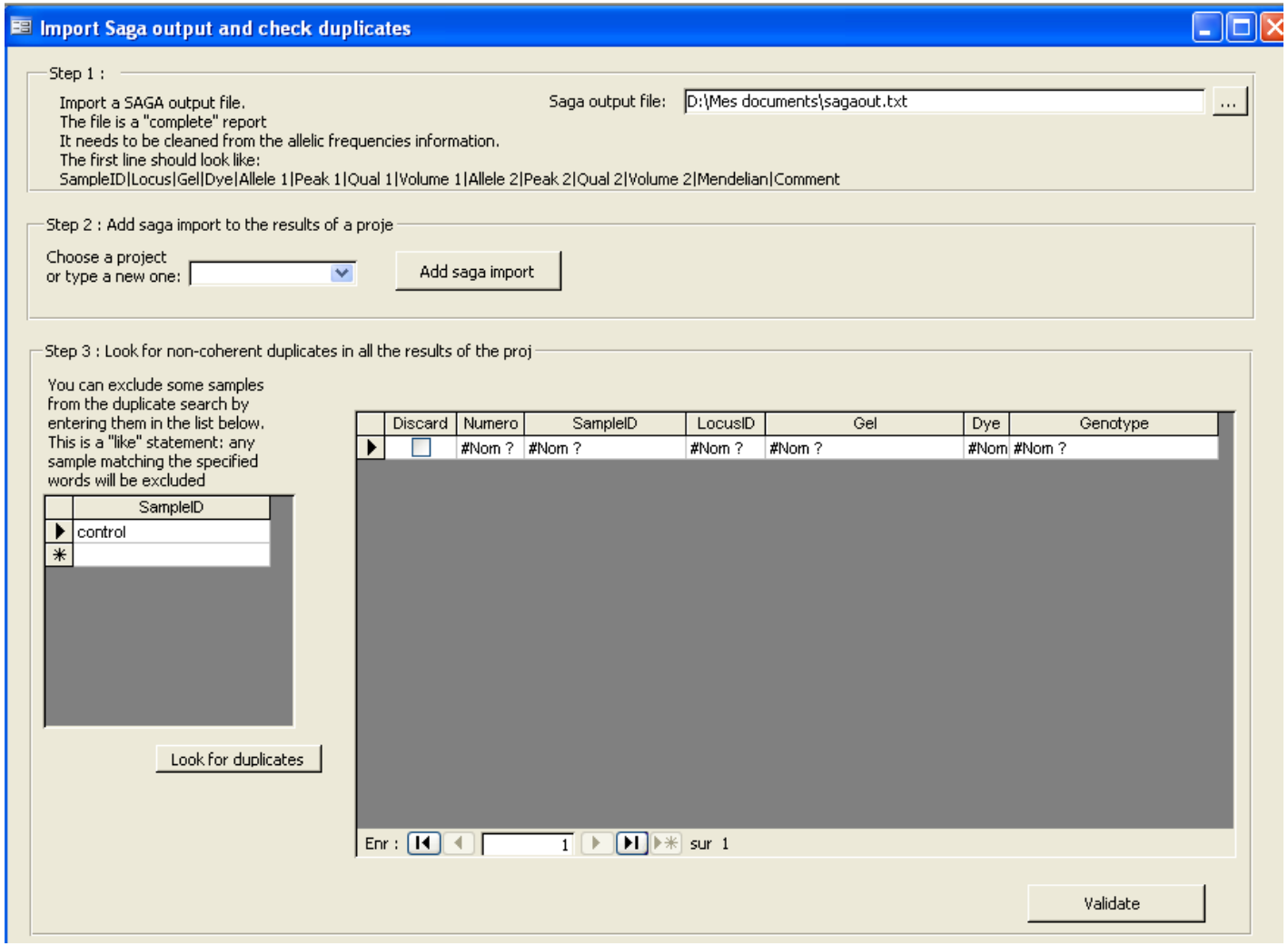
SAGAcity – Import data and check for duplicates

1- Choose datafile exported from SAGA

2- Import it into a project

3- Check for non-coherent duplicates
User control

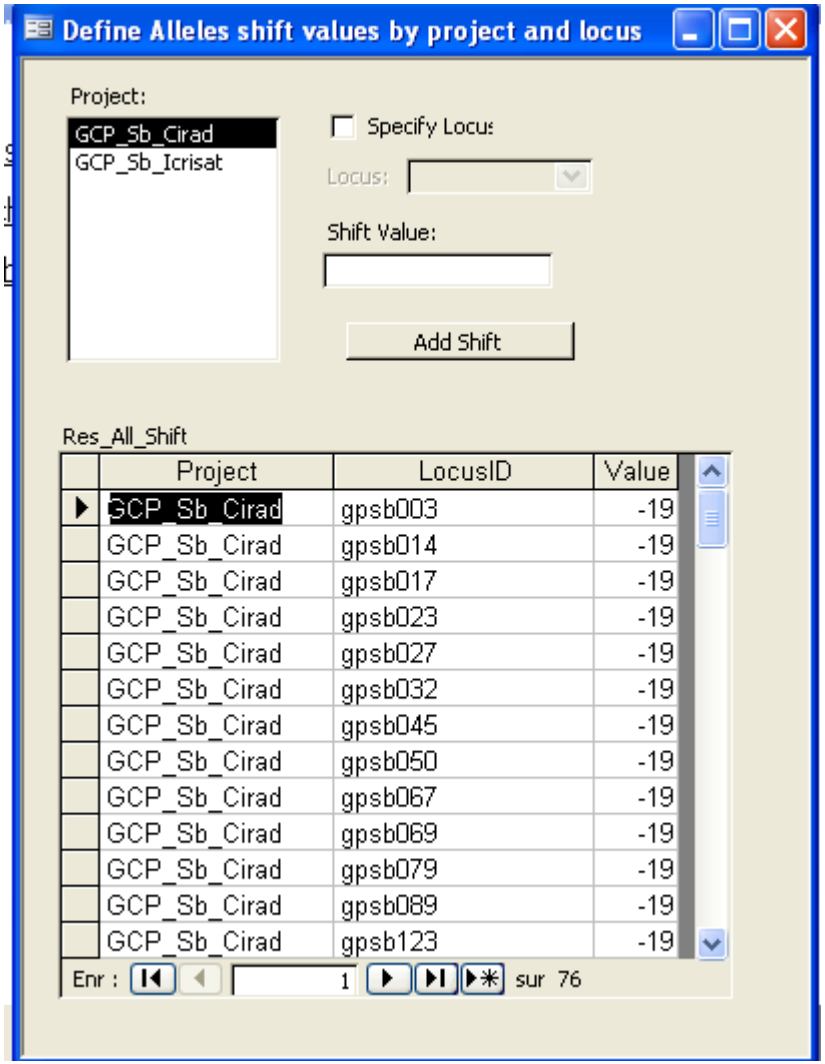
4- Validate the data



SAGAcity – Merging data – Shifting values

According to control values, you can shift all data

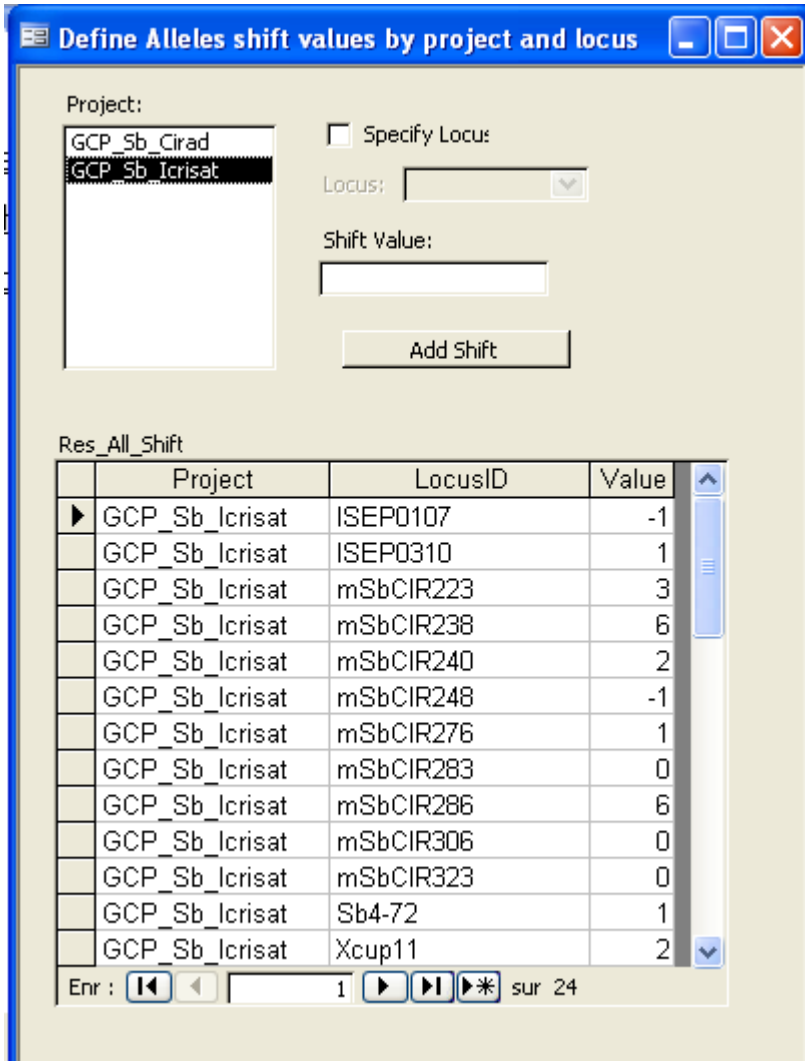
- 19 bp for the M13tail
- shift between labs



SAGAcity – Merging data – Shifting values

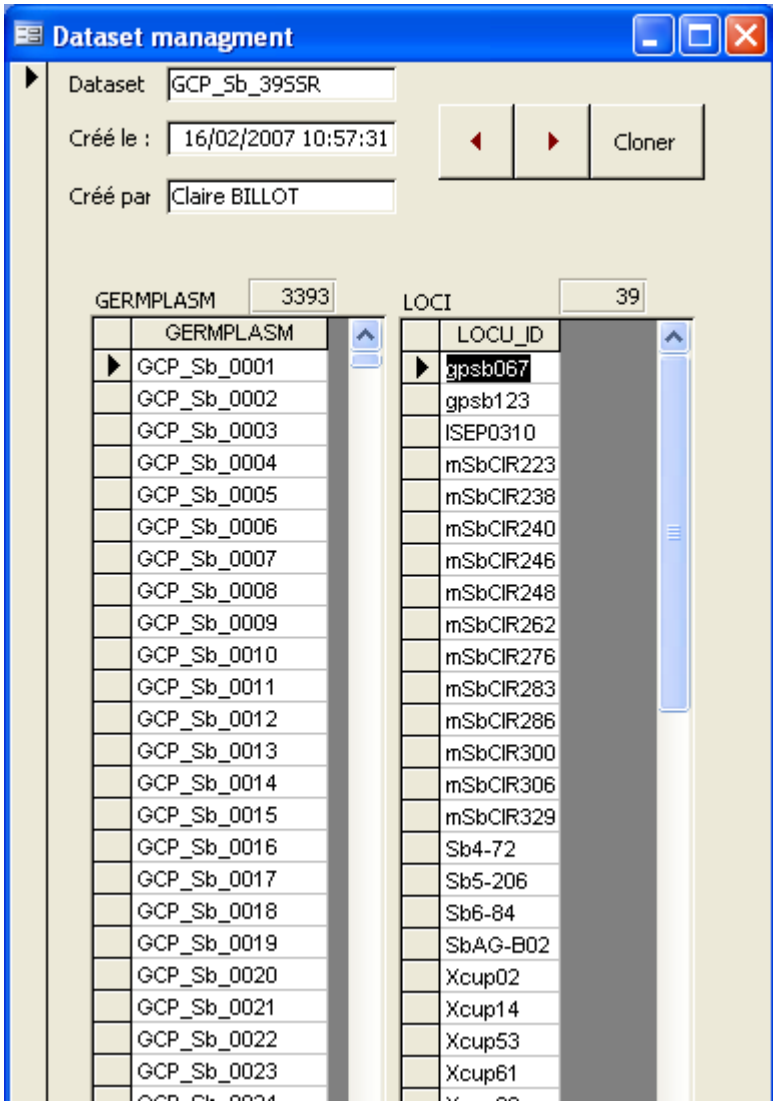
According to control values, you can shift the entire data

- 19 bp for the M13tail
- shift between labs

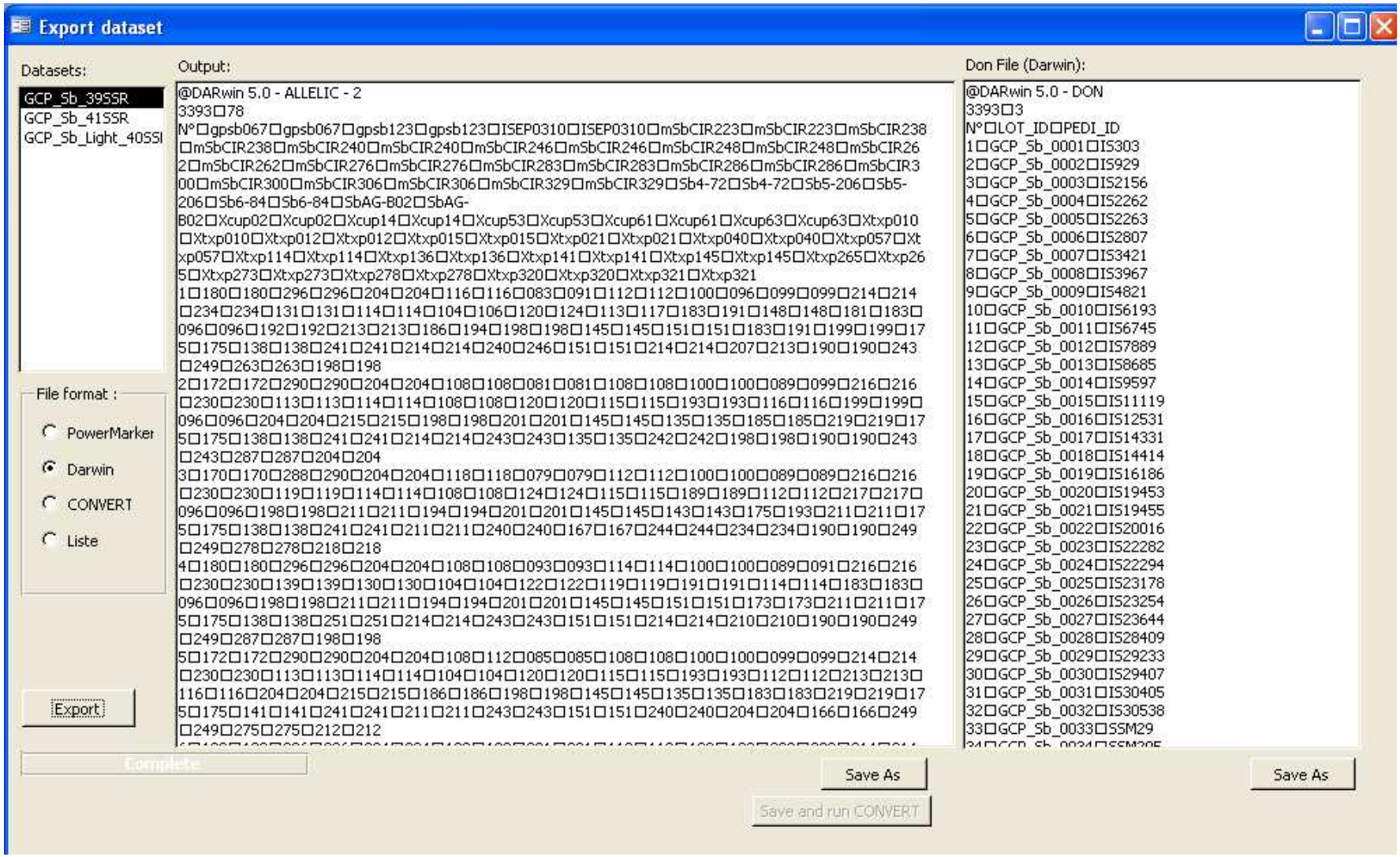


SAGAcity – Defining dataset

- Create a dataset
- Choose the germplasm list
- Choose the locus list



SAGAcity – Exporting dataset



That's it !!!