ISTA
Seed Quality Assurance

ISTA Secretariat
Richtiarkade 18, 8304 Wallisellen, Switzerland
Phone: +41 44 838 60 00 | Fax: +41 44 838 60 01
Email: ista.office@ista.ch
www.seedtest.org

Document OGM22-06

Method Validation Reports on Rules Proposals for the International Rules for Seed Testing 2023 Edition

Validation of a germination method for *Anethum graveolens* using temperature 20°C.
to support C.5.1. Page 2

Validation of a new filtration method for ISTA Method 7-031
to support C.7.1 Page 10

Validation of a new DNA-based method for testing *Pisum* varieties
to support C.8.1 Page 14

Validation of a new DNA-based method for testing *Avena sativa* varieties
to support C.8.2 Page 20

Appendices
Supporting information for validation reports Page 26

Germination Committee Technical report: Validation of a germination method for *Anethum graveolens* using temperature 20°C.

Robineau Lucie, Blouin Valerie and Ducournau Sylvie. GEVES-SNES, France.

Summary

The study was conducted to support the introduction of a new temperature regime (20°C) for the germination of *Anethum graveolens* in the ISTA Rules.

Five laboratories analyzed three lots of *Anethum graveolens*. The temperatures 20°C, 20<=>30°C, and 10<=>30°C were compared, combined with Top of Paper (TP), and Between Paper (BP). Statistical analyses showed that the germination tests performed at 20°C gave results with good repeatability, and results with equal to and even better levels of reproducibility compared with the existing methods for this species in the ISTA Rules. It is therefore suggested that temperature 20°C becomes an additional prescribed temperature for the germination of *Anethum graveolens* in the ISTA Rules.

Introduction

Germination methods currently prescribed in the ISTA Rules for *Anethum graveolens* are Between Paper (BP) and Top of Paper (TP) and the temperatures are 20<=>30°C and 10<=>30°C.

In the National Seed Testing Station in France, germination tests on *Anethum graveolens* are incubated at these different temperatures, but if seed pathogens are present, disease development is generally lower when tested at 20°C, generating a decreased number of abnormal decayed seedlings. After an experiment carried out in the laboratory to compare the existing ISTA methods at 20<=>30°C and 10<=>30°C with a constant temperature at 20°C, the rate of abnormal seedlings decreases considerably. The interest in adding this new temperature is important because it allows a more accurate evaluation of the germination capacity, with consequently a better evaluation of the quality of the seed lot. The introduction of temperature 20°C as a new temperature for *Anethum graveolens* was proposed and approved by the members of the Germination Committee.

A peer validation study was organized to compare the germination results of *Anethum graveolens* obtained with all the current prescribed media and temperatures and the results obtained in the same media with the new temperature.

Materials and Methods

Seed material

Anethum graveolens seed lots were supplied by different seed companies. The three seed lots of *Anethum graveolens* selected for the study were not chemically treated and each was a different variety. Seed lots were tested before starting the study, they did not have any dormancy issues and the germination quality was between 75% and 85% normal seedlings, corresponding to commercial quality standards.

Participant laboratories

A total of five laboratories participated in the *Anethum graveolens* ISTA validation study. The laboratories were located in five countries: Germany, Israel, Italy, the Netherlands, and France.

Germination methods

All the participants compared temperature of 20°C with current ISTA prescribed temperatures. The participants were requested to use the two types of media that are Between paper (BP) and Top of paper (TP). All participants carried out the testing at three temperatures (20°C, 20<=>30°C and 10<=> 30°C). All participants used light during testing, with light varying between 8 to 12 hours.

The duration of the tests was 21days (end of the ISTA test) for each participant. The intermediate counts were done after 7, 14 days and some laboratories added an intermediate count after 10 days.

Table 1 includes the details of the germination methods used by each participant.

All combinations of media, temperature, and seed lots were tested using four replicates of 100 seeds.

The seedling evaluation was based on the seedling group A-2-1-1-1 (ISTA Handbook on Seedling Evaluation, 3rd Edition). The participants provided a description of the abnormal seedlings found in the tests.

Table 1: Germination methods used by each participating laboratory.

Laboratory	Substrate	Temperature (°C)	Light (h)	Final count (days)
1	BP TP	20, 20<=>30, 10<=>30 20, 20<=>30, 10<=>30	8 8	21 21
2	BP TP	20, 20<=>30, 10<=>30 20, 20<=>30, 10<=>30	12 12	21 21
3	BP TP	20, 20<=>30, 10<=>30 20, 20<=>30, 10<=>30	8 8	21 21
4	BP TP	20, 20<=>30, 10<=>30 20, 20<=>30, 10<=>30	8 8	21 21
5	BP TP	20, 20<=>30, 10<=>30 20, 20<=>30, 10<=>30	8 8	21 21

Statistical analyses

Statistical analyses were performed using 'ISTAgermMV', the tool developed by the ISTA Statistics Committee. Boxplots (per lot, per method, per method x lot, and per laboratory), data checking and the repeatability/reproducibility results were generated from this statistical tool.

Results and Discussion

Germination results by seed lot

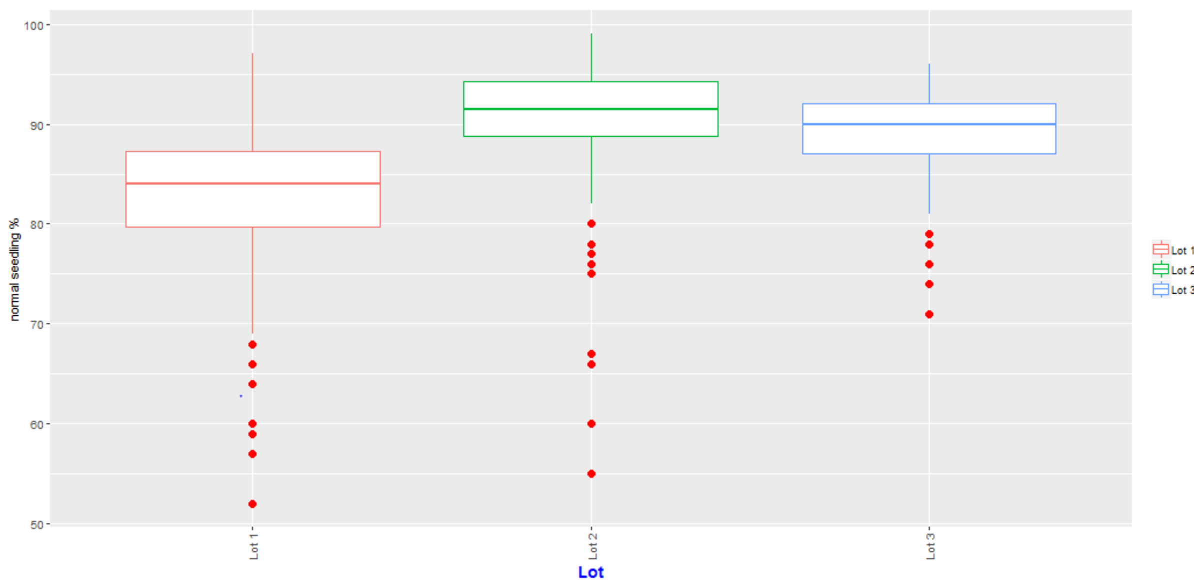

Figure 1: Boxplots for the three seed lots grouped across methods and laboratories.

Figure 1 shows the percentages of normal seedlings obtained for the 3 seed lots, by all the laboratories using all the different methods. Boxplots in the figure show the distribution of the data around the median value. In terms of average results (different from the median values) in ascending order of germination%: seed lot 1 obtained 86.20 % normal seedlings, lot 2 had 87.91% and lot 3 had 88.73%.

Germination results by laboratory

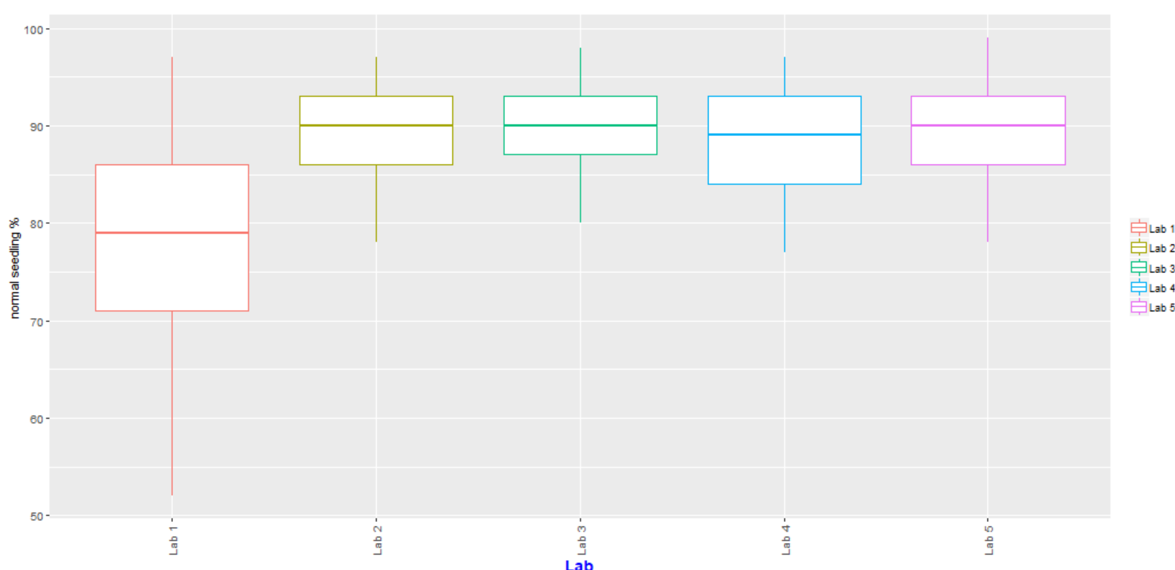

Figure 2: Boxplots for the 5 laboratories grouped across seed lots and methods.

Figure 2 shows the percentages of normal seedlings obtained by all the laboratories, on all seed lots and methods. In terms of laboratory results, lab 1 obtained the lowest germination results (77.4% overall average) and the 4 other participating laboratories obtained better results, all very close to each other, ranging from 88.68 to 89.97%.

Germination results by method

a) For all the participant laboratories

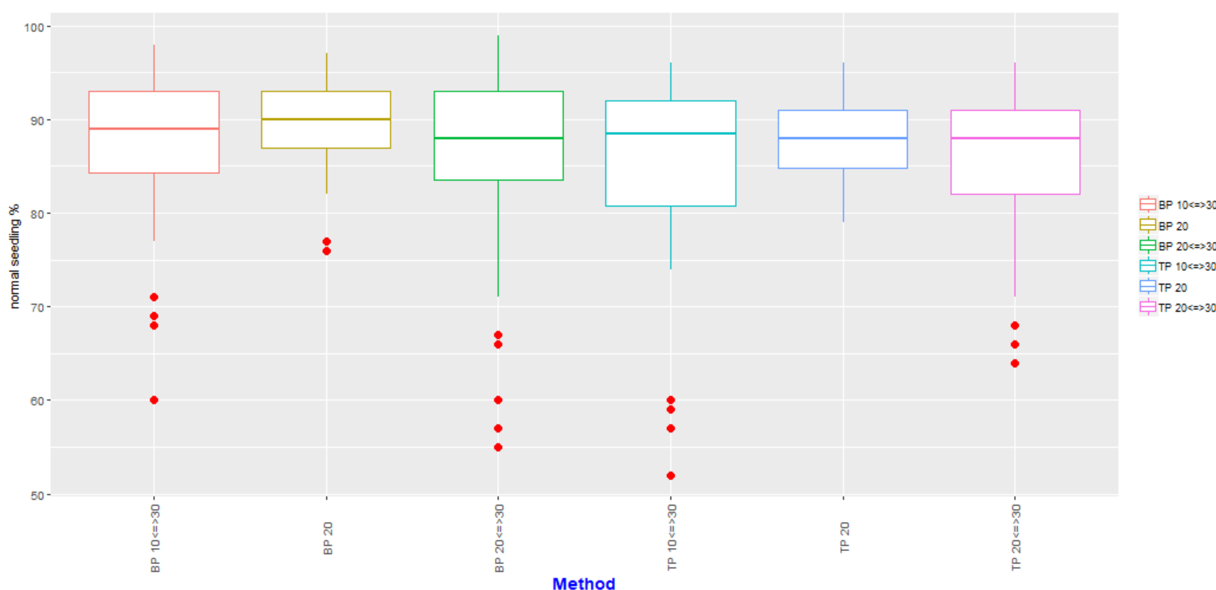

Figure 3a: Boxplots for the 6 methods (media x temperatures) grouped across seed lots and laboratories.

Figure 3a shows the percentages of normal seedlings obtained with the different methods used by all laboratories on the three seed lots. The figure shows that results are more homogeneous, with much less outliers, with the 20°C temperature. The results of the median values are not different depending on the temperatures. When looking at the average results per temperature (10<=>30°C 86%, 20<=>30°C 85.5% and 20°C 89%), the method BP 20°C gives better results than the others.

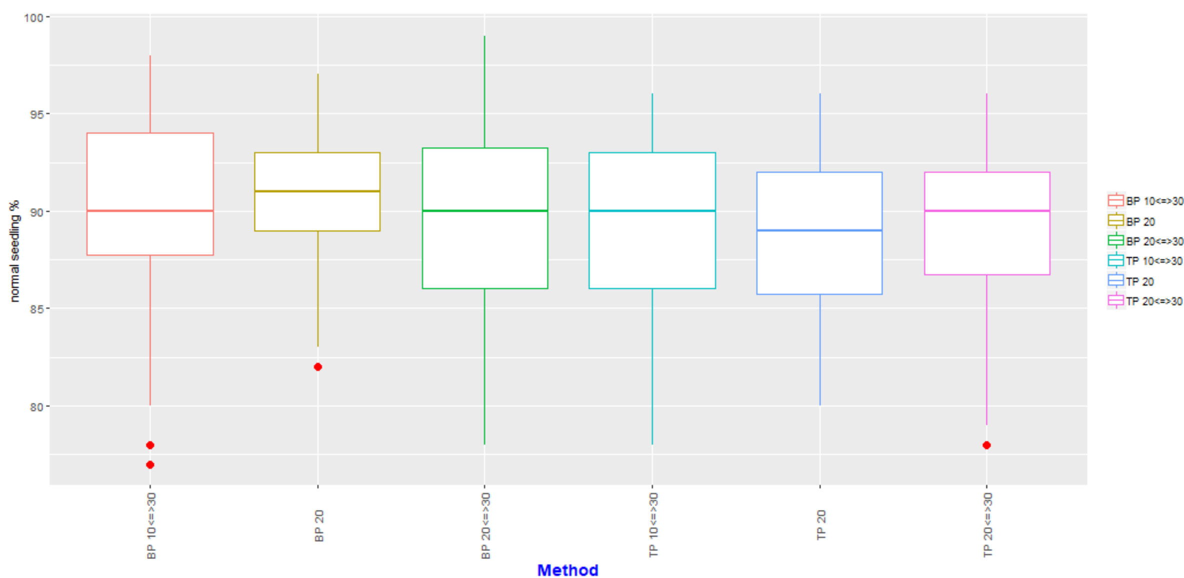

Figure 3b: Boxplots for the 6 methods (media x temperatures) grouped across seed lots and laboratories (without laboratory 1)

Figure 3b shows the percentages of normal seedlings obtained with the different methods used by 4 laboratories on the three seed lots, without lab 1 which obtained lower results than the other labs. The median results or the average results (10<=>30°C 89%, 20<=>30°C 89% and 20°C 90%) do not show differences between the results obtained with the 3 temperatures.

Among the 5 laboratories, 4 noticed seedlings with primary infection on the whole seedling or on the root system. Figure 4 shows the average percentage of abnormal seedlings including decayed seedlings, reported by these 4 laboratories, on the 3 samples tested. It appears clearly that methods using 20°C show less abnormal seedlings (with less infected seedlings) than methods using 10 < = > 30°C and 20 < = > 30°C.

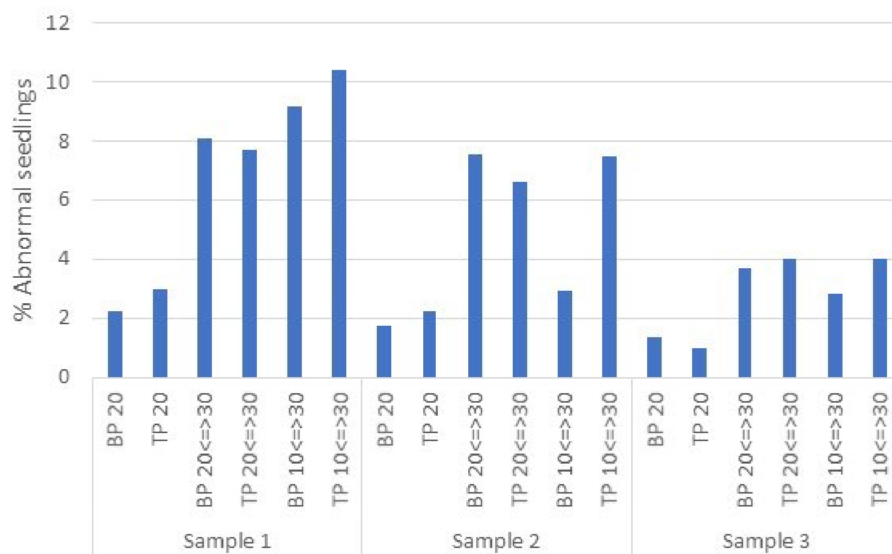

Figure 4: Average percentage of abnormal seedlings reported by the 4 laboratories having reported infected seedlings during their tests (decayed seedlings 00/09 or primary root decayed as a result of primary infection (11/12).

Results of data checking

Data checking of the normal germination percentages was performed according to the ISTA rules by computing tolerances for germination test replicates.

No results were out of tolerance.

Repeatability and reproducibility of the results for the different germination methods

a) For all the participants laboratories

Table 2: Repeatability and reproducibility of the different methods compared for the germination of *Anethum graveolens*

Method	Mean	s_repeatability	disp	s_Reproducibility	s_Lab	s_LotxLab
BP 10<=>30	87.00	3.64	1.09	6.35	4.41	2.75
BP 20	90.00	3.02	0.99	4.57	2.95	1.74
BP 20<=>30	85.00	4.12	1.16	10.91	9.05	4.49
TP 10<=>30	85.00	2.89	0.82	9.32	8.21	3.32
TP 20	88.00	2.93	0.90	3.45	0.00	1.83
TP 20<=>30	86.00	3.59	1.04	7.15	5.68	

b) For the participants without laboratory 1

Table 3: Repeatability and reproducibility of the different methods compared for the germination of *Anethum graveolens*

Method	Mean	s_repeatability	disp	s_Reproducibility	s_Lab	s_LotxLab
BP 10<=>30	89.00	3.29	1.06	3.40	0.00	0.88
BP 20	91.00	2.77	0.95	3.72	2.49	0.00
BP 20<=>30	89.00	3.16	1.03	3.50	0.00	1.50
TP 10<=>30	89.00	2.88	0.92	3.18	1.15	0.70
TP 20	89.00	2.86	0.90	3.12	0.00	1.25
TP 20<=>30	89.00	2.93	0.93	3.87	0.00	2.54

In Tables 2 and 3, summarized by the method (media x temperature), s.repeatability and s.reproducibility corresponds to the repeatability and reproducibility standard deviations. Disp. is the dispersion factor; it applies to the results of repeatability and indicates an over dispersion when its value is greater than 1.
The dispersion factors are around 1 and half of them are less than 1. The values are similar among methods, especially when lab 1 is excluded. Looking specifically at the dispersion factor for the 20°C methods, the values range from 0.90 à to 0.99, indicating good repeatability of the results obtained with these methods in the participating laboratories.

With or without the results from lab 1 (see tables 2 and 3), the germination results are always better using the method BP 20°C. Including the results from lab 1, 20°C with TP or BP gave the better results, but without lab 1, TP 20°C gave lower results.

The values of the standard deviation of reproducibility are different depending on the methods. The lowest values and therefore the best reproducibility is for the method using TP 20°C and BP 20°C when lab 1 is included. When the results of lab 1 are not included, the reproducibility is not very different depending on the method (all the standard deviation values are between 3 and 4). On the other hand, the standard deviation values of reproducibility are higher for the methods using alternating temperatures 10<=>30°C whatever the substrate used, and even more pronounced with the methods using alternating temperature 20<=>30°C.
Germination methods with alternating temperatures give the lowest performance in terms of reproducibility and average germination results (except for TP 20°C without lab 1). This is probably due to the presence of seed health problems in *Anethum graveolens*, with the alternating warm temperatures tending to increase the phenomenon. Results of abnormal seedlings reported by laboratories having found disease symptoms in their tests, give support to this hypothesis.
As a consequence of this study, it is proposed to include 20°C as an additional prescribed temperature for the germination of *Anethum graveolens*.

General conclusion

This study clearly shows that germination of *Anethum graveolens* seeds with a constant temperature of 20°C gives germination results equal to or better than the existing germination methods currently prescribed in the ISTA Rules for this species. These results therefore support the introduction of 20°C as the germination temperature in Table 5A of the Germination chapter of the ISTA Rules.

It is therefore proposed to change the germination methods prescribed in Table 5A for *Anethum graveolens* such as: TP; BP – 20 <=> 30 ; 10 <=> 30 ; 20°C – 21 days – Prechill.

Acknowledgements

Thanks to Enza Zaden, Vivadour and Alliance seeds companies for supplying the seed lots of *Anethum graveolens* for this germination validation study.

Thanks to the labs that participated in this study, including Landwirtschafliches Technologie Zentrum, Karlsruhe, Germany; Official Seed Testing laboratory - Volcani Center, Israel; CREA Consiglio Per La Ricerca Agraria, Italy; Naktuimbouw laboratories, The Netherlands; GEVES – SNES, France.

Thanks to the ISTA Technical reviewers Hortense Faucher (Corteva -France) and Sarah Dammen (SGS South Dakota – USA) the members of the ISTA Germination Committee.

Thanks to Jean-Louis Laffont, Statistical Committee chair, for the ISTAgermMV stat program to run the data and the support he gave for interpretation.

References

1. International Seed Testing Association. 2021 Edition. ISTA International Rules for Seed Testing. Bassersdorf, Switzerland.
2. International Seed Testing Association. 2018 Edition. ISTA Handbook on Seedling Evaluation, Fourth Edition. Bassersdorf, Switzerland.

Validation of new container for the ISTA method 7-031

Authors

Serandat Isabelle, Beduneau Helena, Avrillon Myriam and Grimault Valerie. GEVES-SNES, France.

Background

At OGM 2017, the ISTA membership voted to approve proposal C.7.2. (Addition of new method. 7-031: Filtration method for detection of *Ditylenchus dipsaci* in *Medicago sativa*; *D. dipsaci* and *D. gigas* in *Vicia faba* seed). In the proposal, the use of a 250µm sieve covered with soft filter paper was recommended. The full validation study for the original study may be found in the OGM 17 published documents. The current study compared the original filtration recommendation with an option for an equivalent nematode-permeable container for methods in 7-031.

Materials and methods

Geves validated a new container (figures 1-2-3) by comparing it to the current 250µm sieve recommendation (figures 4-5-6). This new container is a non-woven plant growth bag 16x20cm (KINGLAKE brand). Use of the new container for comparison purposes, was the only change in the procedure as written in 7-031. All other materials and methods used for sieving and confirmation followed current ISTA procedures.

Figure 1 : New container before soaking

Figure 2 : New container soaking

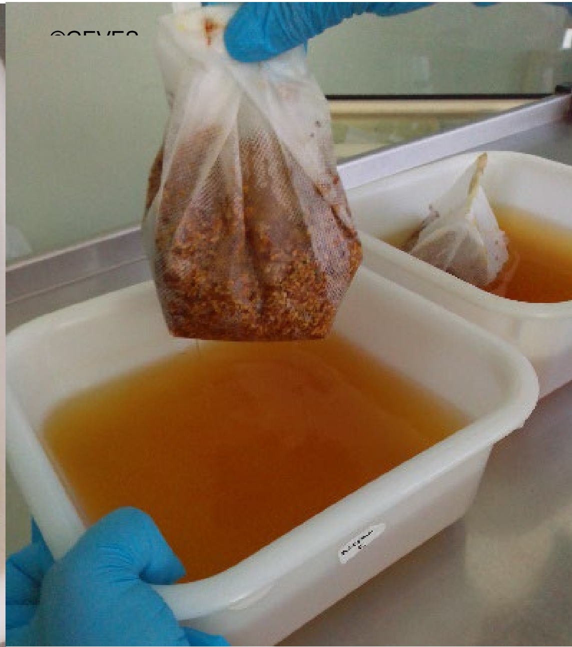

Figure 3 : New container removal

Figure 4 : 250µm sieve

Figure 5 : 250µm sieve soaking

Figure 6 : 250µm sieve removal

Participant Laboratories

Geves, 25 rue Georges Morel, CS 90024, 49071 Beaucouze, France

Summary of results

Analytical sensitivity: 2 tests were caried out.

Test 1

Modality	Spiking
A	3 *Ditylenchus dipsaci*
B	10 *Ditylenchus dipsaci*
C	3 *Ditylenchus dipsaci* + saprophagous
D	10 *Ditylenchus dipsaci* + saprophagous
E	0 *Ditylenchus dipsaci*

Table 1 : Analytical sensitivity results of test 1

				Test 1		Test 2	
Modality	**N° sample**	**Nb *Dd***	**Nb *Dd* minimum expected**	**Nb *Dd* obtained new container**	**Nb *Dd* obtained 250µ sieve**	**Nb *Dd* obtained new container**	**Nb *Dd* obtained 250µ sieve**
	1	3	1	3	1	3	1
	2	3	1	3	2	3	2
A	3	3	1	2	3	2	3
	4	10	2	8	9	7	9
	5	10	2	10	7	7	7
B	6	10	2	9	7	6	7
	7	3+sapro	1	3	3	3	3
	8	3+sapro	1	2	3	3	3
C	9	3+sapro	1	3	3	2	3
	10	10+sapro	2	6	10	8	10
	11	10+sapro	2	8	6	6	6
D	12	10+sapro	2	6	10	8	10
	13	0	0	0	0	0	0
	14	0	0	0	0	0	0
E	15	0	0	0	0	0	0

Test 2

10 spiked samples and one negative sample were tested with the new container:

Table 2: Analytical sensitivity results of test 2

N° sample	**Nb *Dd***	**Nb *Dd* minimum expected**	**Nb *Dd* obtained new container**
1	5	1	4
2	5	1	4
3	5	1	4

4	5	1	4
5	5	1	5
6	5	1	5
7	5	1	4
8	5	1	4
9	5	1	5
10	5	1	4
11	0	0	0

Result: The minimum expected was always obtained.

Diagnostic sensitivity and specificity (accuracy)/repeatability-reproducibility

Three samples with different levels of contamination were analysed with the new container in 3 replicates and by two different analysts.

The 250µm sieve was used as a reference.

- Healthy lot: expected result negative
- Naturally infected lot (2 to >100 nematodes): expected result positive
- Artificially contaminated lot (10 nematodes): expected result positive

Table 3 : Statistical analysis of accuracy

	Expected result +	Expected result -	sensitivity	specificity	accuracy
Obtained result +	6	0	100.00	100.00	100.00
Obtained result -	0	3			

The conclusion of the statistical analysis is that given these results, there is enough evidence for validating the method.

Validation of a new method for a "DNA based test on PEAS"

Authors

Dr. Marie-José Côté – Canadian Food Inspection Agency, Canada. CT leader. Former Variety Committee member.

Dr. Marie-Claude Gagnon – Canadian Food Inspection Agency, Canada. Variety Committee Member.

Dr. Ana Laura Vicario – Instituto Nacional de Semillas, Argentina. Variety Committee Chair.

Background

Traditionally, ISTA's standardized procedures for the determination of variety verification have been based on the examination of seeds, seedlings or plants in a laboratory, glasshouse, growth chamber or field plot, to assess morphology (grow-out tests), specific substances (biochemical methods) or protein characteristics (protein-based methods).

DNA-based approaches are very useful tools for variety verification and for assessment of purity. In comparison to traditional variety verification methods, DNA-based techniques may reveal more polymorphism thus allowing greater resolution among varieties. DNA-based techniques are also independent of environmental conditions or developmental stages. In 2017, a wheat DNA-based test was included for the first time in Chapter 8 of ISTA Rules. Following that, maize was also included as a new method in the Rules.

Now we propose the inclusion of a new method for peas variety verification by means of microsatellites markers.

Introduction

To initiate the process for the incorporation of peas DNA-based markers into the Rules, Dr Marie-José Côté from the Variety Committee organized two comparative tests (CTs) for peas with participation of laboratories from several countries from around the world over a period of two years.

The objectives of the CTs carried out are summarized below:

- The aim of CT1 was to compare results among participant laboratories and evaluate the possibility of obtaining the same SSR profiles and same allele sizes using different reagents, equipment and working protocols. Varieties and SSR markers were the same for all participant laboratories.

The expected result of CT1 was to obtain comparable results among laboratories.

- The aim of CT2 was also to compare results among participant laboratories and evaluate the possibility of obtaining the same SSR profile and same allele sizes when using different reagents, equipment and working protocols. Varieties and SSR markers were the same for all participant laboratories.

However, during CT2, the applicability of the method was evaluated by extending the range of varieties tested compared to CT1. Also, reproducibility of the markers was tested by comparing results among more laboratories than for CT1. Repeatability was tested with one single lab that ran both CTs.

Materials and methods

Samples:

For CT1, 8 varieties were analyzed (1 from Serbia, 3 from U.S. and 4 from Canada) using 11 SSR markers with long history of use in Canada, obtained from Loridon et al. (2005). Each participant received 10 individual 1/2 seed per variety and a subsample of a pool of all the other halves. To minimize inter-laboratory variation, 2 Standards (1/2 seeds) with known genotypes have been included in this panel. Together, these Standard samples are generating a minimum of 2 alleles per marker (except one) and therefore will help calibrate the system. Samples were provided in sealed tubes labelled with the name or code of the variety.

For CT2, 24 varieties were analyzed (4 from France, 8 from Canada, 4 from South Africa, 4 from Spain and 4 from Italy) using the 11 SSR markers tested in CT1. Each participant received 2 tubes with subsamples of a pool of 20 crushed seeds per variety. To minimize inter-laboratory variation, 3 Standards with known genotypes have been included in this panel in duplicates (see CT2 results excel sheet – pea for the genotypes). Together, these Standard samples were generating a minimum of 2 alleles per marker and therefore helped calibrate the system for allele calls. Samples were provided in sealed tubes labelled with the name or code of the variety. Duplicates were provided to all participants as a backup. If any participating laboratory ran out of a sample, they could contact CT leader to request more material.

Technical information:

1. Any protocols can be used to perform DNA extraction or PCR amplification
2. Amplicon detection must be carried out using Sequencing type detection systems such as LiCor or ABI capillary sequencers or the equivalent
3. Markers can be run individually or multiplexed.

DNA extraction	Qiagen DNeasy Plant extraction kit
Primer labelling and detection system	No tailed forward primers labelled with fluorophores compatible with ABI DNA fragment analysers running POP7
Primer multiplexing	Each marker can be run individually or multiplexed for the CT. If multiplexed, it is recommended to use PCR mixes that are designed for multiplexing such as Type-it microsatellite PCR kit from Qiagen. However, it may be still required to optimize the primer concentration within a multiplex
PCR condition for thermocycler using Type-it microsatellite PCR kit from Qiagen	Hold: 5'@ 95°C (Initial denaturation conditions according to manufacturer of Taq polymerase used) 2. 28 cycles: 30s @ 95°C, 90s@ 60°C, 30s@72°C 3. Hold: 30min@ 60°C 4. Hold: forever @ 4°C
Allele calls	Alleles are called according to their sizes in base pairs. However, since every detection system and chemical used differ from one lab to another, this may alter the migration of the PCR products. To minimise the interlaboratory variation, we have included in the panel, samples with known genotype (Standards 1, 2, 3). Together these samples are to generate a minimum of 2 alleles per markers to help calibrating your system.
Additional general information specific to the crop and the marker system	"Null" allele (No PCR amplification) are known to occur in peas. NULL is called when repeated attempts to amplify a specific marker fail to produce a detected allele for specimens of the same variety, while the others markers work well. It would be advisable to confirm the absence of amplification for these particular samples to rule out chemical and enzyme malfunction.

SSR Primer Sets

SSR	Forward 5' to 3'	Reverse 5' to 3'	Size range of PCR product in base pairs	Final concentration in PCR mix *
A9	GTGCAGAAGCATTTGTTCAGAT	CCCACATATATTTGGTTGGTCA	340-420	[3.75 µM]
AA67	CCCATGTGAAATTCTCTTGAAGA	GCATTTCACTTGATGAAATTTCG	370-425 **	[1.25 µM]
AA135	CCGTTACACATCATTAAGATG	TCCATATCCAGATTAGTCAGA	360-370 **	[3.0 µM]
AA205	TACGCAATCATAGAGTTTGGAA	AATCAAGTCAATGAAACAAGCA	220-250 **	[1.25 µM]
AA285	TCGCCTAATCTAGATGAGAATA	CTTAACATTTTAGGTCTTGGAG	230-260	[2.5 µM]
AA355	AGAAAAATTCTAGCATGATACTG	GGAAATATAACCTCAATAACACA	180-250	[3.75 µM]
AB72	ATCTCATGTTCAACTTGCAACCTTTA	TTCAAAACACGCAAGTTTTCTGA	250-290	[2.0 µM]
AD59	TTGGAGAATGTCTTCTCTTTAG	GTATATTTTCACTCAGAGGCAC	310-340 **	[5.0 µM]
AD73	CAGCTGGATTCAATCATTGGTG	ATGAGTAATCCGACGATGCCTT	220-280	[2.5 µM]
AD270	CTCATCTGATGCGTTGGATTAG	AGGTTGGATTTGTTGTTTGTTG	250-320 **	[1.25 µM]
D23	ATGGTTGTCCCAGGATAGATAA	GAAAACATTGGAGAGTGGAGTA	170-200	[3.0 µM]

* May vary depending on reagents - this is used with 1ng of DNA per µl of PCR reaction

** Marker that can show Null allele: see definition in the technical information sheet of this document

Varieties list:

Variety name/code	Country	Sample type	Number of subsample
Standard 1	N/A	Pool	2
Standard 2	N/A	Pool	2

Standard 3	N/A	Pool	2
CDC Winfield	Canada	Pool	2
Miami	Canada	Pool	2
Karina	France	Pool	2
Merveille de Kelvedon	France	Pool	2
Normand	France	Pool	2
Petit Provencal	France	Pool	2
AAC Carver	Canada	Pool	2
AAC Lancombe	Canada	Pool	2
Abarth	Canada	Pool	2
Agassiz	Canada	Pool	2
CDC Inca	Canada	Pool	2
CDC Spectrum	Canada	Pool	2
Arvika	South Africa	Pool	2
Astronaute	South Africa	Pool	2
Emperor	South Africa	Pool	2
Greenfeast	South Africa	Pool	2
ESP-P-1	Spain	Pool	2
ESP-P-2	Spain	Pool	2
ESP-P-3	Spain	Pool	2
ESP-P-4	Spain	Pool	2
IT5	Italy	Pool	2
IT6	Italy	Pool	2
IT7	Italy	Pool	2
IT8	Italy	Pool	2

Equipment, chemicals and procedure

Inclusion of DNA-based methods into the Rules is semi-performance based. Laboratories were provided with guidelines for running the SSR prescribed, but finally the specific procedure was up to the participating laboratory.

Evaluation and reporting of results

Results were reported in an Excel sheet indicating laboratory number, variety name, SSR name, and allele sizes.

The data analysis from CT1 aimed to evaluate if the marker panel was reproducible among laboratories and thus suitable for be kept for CT2 and eventually for the Rules proposal. This evaluation was carried out by the crop leader and consisted of verifying if markers gave the same allele's pattern across laboratories (even using different equipment and reagents). The evaluation of allelic profiles gave comparable results among laboratories; thus, the selected marker panel was deemed appropriate for CT2.

For CT2, the group leader compiled the results and prepared an Excel file with allele sizes and binary data. Binary data was sent to STACOM chair for their analysis (Appendix 1).

Participating Laboratories

CT1 Participating Laboratories

- Ksenija Taški-Ajduković – Institute of Field and Vegetable Crop (Serbia)
- Jeffrey Prischmann – North Dakota State Seed (U.S.A.)
- Marie-Claude Gagnon – Canadian Food Inspection Agency (Canada)
- Alex Reid – Science and Advice for Scottish Agriculture (UK)
- Daniel Perry – Canadian Grain Commission, Grain Research Laboratory (Canada)

Three laboratories sent a data package.

CT2 Participating Laboratories

- Tertia Erasmus; SciCorp Laboratories (South Africa)
- Stephanie Guillet; Eurofins (France)
- Nicole Calliou; SGS BioVision (Canada)
- Luz María Paz; INIA (Spain)
- Daniel Perry; CGC (Canada)
- Marie-Claude Gagnon; CFIA (Canada)
- Chiara Delogu; CREA (Italy)
- Kim Kenward; 2020 Seed lab (Canada)
- Anne Bernole; BioGEVES (France)

Seven laboratories sent a data package.

Statistical analysis

Overall percentage agreements (p_a) and Cohen's kappas have been computed for all the possible laboratory pairs, considering as units either the marker alleles or the varieties. The computations have been performed with the R *irr* package (Gamer *et al*., 2012) which includes functions for computing various coefficients of reliability of agreement.

The overall agreement percentage for scoring varieties across laboratories is above 90% for all the varieties except one (variety 11) when comparing laboratory 2 with laboratory 5 (p_a = 89%). Most of Cohen's kappa values are above 0.6 for scoring varieties and only few Cohen's kappa values are below 0.6 for scoring alleles.

The conclusion of the statistical analysis is that given these results, there is enough evidence for validating the method. For statistical report details see Appendix: 1 and 2

Final comments and conclusions

After running two comparative tests for peas with the participation of 10 laboratories around the world, a far used SSR pea panel and varieties representing the variability worldwide, the statistical analysis done by STACOM concludes that there was enough evidence for validating the method.

Given the work carried out and the STACOM conclusion, the Variety Committee presents this validation report for considering the inclusion of the pea SSR marker panel in Rules Chapter 8.

Reference documents

- See Appendices: 1, 2, and 3
- Loridon K., McPhee K., Morin J., Dubreuil P., Pilet-Nayel M. L., Aubert G., Rameau C., Baranger A., Coyne C., Lejene-Hénaut I. and Burstin J. (2005) Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.). Theor Appl Genet 111: 1022–1031
- Gamer, M., Lemon, J., Fellows, I. and P. Sing (2012). irr: Various coefficients of interrater reliability and agreement. R Package version 0.84. http://CRAN.R-project.org/package=irr.

Validation of a new method for a “DNA based test on OAT”

Authors

Dr. Marie-José Côté – Canadian Food Inspection Agency, Canada. CT leader. Former Variety Committee member.

Dr. Marie-Claude Gagnon – Canadian Food Inspection Agency, Canada. Variety Committee Member.

Dr. Ana Laura Vicario – Instituto Nacional de Semillas, Argentina. Variety Committee Chair.

Background

Traditionally, ISTA’s standardized procedures for the determination of variety verification have been based on the examination of seeds, seedlings or plants in a laboratory, glasshouse, growth chamber or field plot, to assess morphology (grow-out tests), specific substances (biochemical methods) or protein characteristics (protein-based methods).

DNA-based approaches are very useful tools for variety verification and for assessment of purity. In comparison to traditional variety verification methods, DNA-based techniques may reveal more polymorphism thus allowing greater resolution among varieties. DNA-based techniques are also independent of environmental conditions or developmental stages. In 2017, a wheat DNA-based test was included for the first time in Chapter 8 of ISTA Rules. Following that, maize was also included as a new method in the Rules.

Now we propose the inclusion of a new method for oats variety verification by means of microsatellites markers.

Introduction

To initiate the process for the incorporation of oats DNA-based markers into the Rules, Dr Marie-José Côté from the Variety Committee organized two comparative tests (CTs) for oats with participation of laboratories from several countries from around the world over a period of two years.

The objectives of the CTs carried out are summarized below:

- The aim of CT1 was to compare results among participant laboratories and evaluate the possibility of obtaining the same SSR profiles and same allele sizes using different reagents, equipment and working protocols. Varieties and SSR markers were the same for all participant laboratories.

The expected result of CT1 was to obtain comparable results among laboratories.

- The aim of CT2 was also to compare results among participant laboratories and evaluate the possibility of obtaining the same SSR profile and same allele sizes when using different reagents, equipment and working protocols. Varieties and SSR markers were the same for all participant laboratories.

However, during CT2, the applicability of the method was evaluated by extending the range of varieties tested compared to CT1. Also, reproducibility of the markers was tested by comparing results among more laboratories than for CT1. Repeatability was tested with one single laboratory that ran both CTs.

Materials and methods

Samples:

For CT1, 8 varieties were analysed (3 from Austria, 3 from U.S. and 2 from Canada) using 9 SSR markers obtained from Li et al. (2000) and Wight et al. (2010). Each participant received 10 individual crushed seeds per variety and a subsample of a pool of 30-40 seeds. To minimize inter-laboratory variation, 3 Standards (pool of crushed seeds) with known genotypes were included in this panel (see CT results – oat, for the genotypes). Together, these Standard samples were generating a minimum of 2 alleles per marker (except one) and therefore helped to calibrate the system. Samples were provided in sealed tubes labelled with the name or code of the variety.

For CT2, 24 varieties were analysed (9 from Canada, 4 from Austria, 4 from South Africa, 4 from Italy and 3 from Spain) using the 9 SSR markers tested in CT1. Each participant received 2 tubes with subsamples of a pool of 40 crushed seeds per variety. To minimize inter-laboratory variation, 3 Standards with known genotypes were included in this panel in duplicate (see CT results – oat for the genotypes). Together, these Standard samples were generating a minimum of 2 alleles per marker and therefore helped to calibrate the system for allele calls. Samples were provided in sealed tubes labelled with the name or code of the variety. Duplicates were provided to all participants as a backup. However, if any participating laboratory ran out of a sample, they were able to contact CT leader to request more material.

Technical information:

1. Any protocols can be used to perform DNA extraction or PCR amplification.
2. Amplicon detection must be carried out using Sequencing type detection systems such as LiCor gel system or ABI capillary sequencers or the equivalent
3. Markers can be run individually or multiplexed.

DNA extraction	Qiagen DNeasy Plant extraction kit
Primer labelling and detection system	No tailed forward primers are labelled with fluorophores compatible with ABI DNA fragment analysers running POP7
Primer multiplexing	Each marker can be run individually or multiplexed for the CT. If multiplexed it is recommended to use PCR mixes that are designed for multiplexing such as Type-it microsatellite PCR kit from Qiagen. However, it may be still required to optimize the primer concentration within a multiplex
PCR condition for thermocycler for AM markers using Type-it microsatellite PCR kit from Qiagen	Hold: 5'@ 95°C (Initial denaturation conditions according to manufacturer of Taq polymerase used)
	2. 10 cycles: 30s @ 95°C, 90s@ 63°C (-0.5°C/cycle), 30s@72°C
	3. 20 cycles: 30s @ 95°C, 90s@ 58°C, 30s@72°C
	4. Hold: 30'@ 60°C
	5. Hold: forever @ 4°C
PCR condition for thermocycler for MAMA markers using Type-it microsatellite PCR kit from Qiagen	Hold: 5'@ 95°C (Initial denaturation conditions according to manufacturer of Taq polymerase used)
	2. 1 cycle: 1'@ 58°C, 1'@72°C
	3. 29 cycles: 30s @ 95°C, 90s@ 58°C, 30s@72°C
	4. Hold: 30'@ 60°C
	5. Hold: forever @ 4°C
Allele calls	Alleles are called according to their sizes in base pairs. However, since every detection system and chemical used differ from one lab to another, this may alter the migration of the PCR products. To minimise the interlaboratory variation, we have included in the panel, samples with known genotype (Standards 1, 2, 3). Together these samples are to generate a minimum of 2 alleles per markers to help calibrating your system.
	NULL alleles: Occasionally NULL alleles have been observed. NULL is called when repeated attempts to amplify a specific marker fail to produce a detected allele for samples of the same variety, while the others markers work well.

SSR Primer Sets:

SSR	Forward 5' to 3'	Reverse 5' to 3'	Size range of PCR product in base pairs	Final concentration in PCR mix *
AM031	GCAAAGGCCATATGGTGAGAA	CATAGGTTTGCCATTCGTGGT	145-195	[0.18 µM]
AM102	TGGTCAGCAAGCATCACAAT	TGTGCATGCATCTGTGCTTA	206-220	[0.85 µM]
AM112	AGCGGTGTAGGGGAAAGAGT	TTCTTGGTTTAGATGGGAGGA	233-266	[0.35 µM]
MAMA-1	GTGCGCCTCTAACGAAAAAT	CATGCTGGCGAAATCTATCA	162-226	[0.55 µM]
MAMA-3	ATGTTCTCCAATGGGACTGC	ATCGCGATGACTGTGTGA	364-402	[0.18 µM]
MAMA-4	GGAGTGGGCGTTTGACATTA	CAGCTACCGGTTTTCATTCC	290-366	[0.39 µM]
MAMA-5	GGATTGGGACTTCGCATCTA	AACCCTAATTACTGCTCCGTTTC	156-254	[0.30 µM]
MAMA-6	GACTAAATCACACAACCCAACC	GCAGAATCGCGGGAAAGA	167-223	[1.10 µM]
MAMA-7	ATAAATACGCGCCACCACTC	TCCGGTGTGAGTAGGGTAGG	324-362	[0.80 µM]

* May vary depending on reagents - this is used with 1ng of DNA per µl of PCR reaction

Varieties List:

Variety name/code	Country	Sample type	Number of subsample
Standard 1	N/A	Pool	2
Standard 2	N/A	Pool	2
Standard 3	N/A	Pool	2
AC Rebel	Canada	Pool	2
Niagara	Canada	Pool	2
Robert	Canada	Pool	2
A	Austria	Pool	2
B	Austria	Pool	2
C	Austria	Pool	2
D	Austria	Pool	2
AC Stride	Canada	Pool	2
CDC Baler	Canada	Pool	2
CDC Haymaker	Canada	Pool	2
CS Camden	Canada	Pool	2
Juniper Oats	Canada	Pool	2
ORe3542M	Canada	Pool	2
Horsepower	South Africa	Pool	2
Kompasberg	South Africa	Pool	2
Overberg	South Africa	Pool	2
SSH 39W	South Africa	Pool	2
ESP-O-1	Spain	Pool	2
ESP-O-2	Spain	Pool	2

ESP-O-3	Spain	Pool	2
IT1	Italy	Pool	2
IT2	Italy	Pool	2
IT3	Italy	Pool	2
IT4	Italy	Pool	2

Equipment, chemicals and procedure

Inclusion of DNA-based methods into the Rules is semi-performance based. Laboratories were provided with guidelines for running the SSR prescribed, but finally the specific procedure was up to the participating laboratories.

Evaluation and reporting of results

Results were reported in an Excel sheet indicating laboratory number, variety name, SSR name, and allele sizes.

The data analysis from CT1 aimed to evaluate if the marker panel was reproducible among laboratories and thus suitable for be kept for CT2 and eventually for the Rules proposal. This evaluation was carried out by the crop leader and consisted on verifying if markers gave the same allele's pattern across laboratories (even using different equipment and reagents). The evaluation of allelic profiles gave comparable results among laboratories, thus the selected marker panel was deemed appropriate for CT2.

For CT2, the group leader compiled the results and prepared an Excel file with allele sizes and binary data. Binary data was sent to STACOM chair for their analysis (Appendix 4).

Participating Laboratories

CT1 Participating Laboratories

- Marie-Claude Gagnon. Canadian Food Inspection Agency, Ottawa Plant Laboratory. Ottawa, Ontario, Canada.
- Daniel Perry. Canadian Grain Commission, Grain Research Laboratory. Winnipeg, Manitoba, Canada.
- Doris Kaiser/Verena Peterseil. Austrian Institute for Food Safety. Vienna, Austria.
- Jeffrey Prischmann. North Dakota State Seed Department. Fargo, North Dakota, USA.

Three laboratories sent a data package.

CT2 Participating Laboratories

- Tertia Erasmus; SciCorp Laboratories (South Africa)
- Nicole Calliou; SGS BioVision (Canada)
- Edurne Aguiriano Labandibar/Luz María Paz; INIA (Spain)
- Verena Peterseil/Doris Kaiser; AGES (Austria)
- Daniel Perry; CGC (Canada)

- Marie-Claude Gagnon; CFIA (Canada)
- Chiara Delogu; CREA (Italy)
- Kim Kenward; 20/20 Seed Labs (Canada)
- Anne Bernole; BioGEVES (France)

Seven laboratories sent a data package.

Statistical analysis

Overall percentage agreements (p_a) and Cohen's kappas have been computed for all the possible laboratory pairs, considering as units either the marker alleles or the varieties. The computations have been performed with the R irr package (Gamer et al., 2012) which includes functions for computing various coefficients of reliability of agreement.

The overall agreement percentage for scoring varieties across laboratories is above 90% for all the varieties except one (variety 16) when comparing laboratory 4 with laboratory 7 (pa = 88.7%). Most of Cohen's kappa values are above 0.8 for scoring varieties and only few Cohen's kappa values are below 0.6 for scoring alleles.

The conclusion of the statistical analysis is that given these results, there is enough evidence for validating the method. For statistical report details see Appendices 4 and 5.

Final comments and conclusions

After running two comparative tests for oats with the participation of 10 laboratories around the world, a far used SSR oat panel and varieties representing the variability worldwide, the statistical analysis done by STACOM concludes that there was enough evidence for validating the method.

Given the work carried out and the STACOM conclusion, the Variety Committee presents this validation report for considering the inclusion of the oat SSR marker panel in Rules Chapter 8.

Reference documents

- See Appendices: 4, 5, and 6
- C.D. Li, B.G. Rossnagel and G.J. Scoles (2000). The development of oat microsatellte markers and their use in identifying relationships among *Avena* species and oat cultivars. Theoretical and Applied Genetics 101:1259–1268.
- C. P. Wight, W, Yan, J. Mitchell Fetch, J. Deyl and N.A. Tinker (2010) A set of new simple sequence repeat and avenin DNA markers suitable for mapping and fingerprinting studies in oat (*Avena spp.*). Crop Science 50:1207–1218.
- Gamer, M., Lemon, J., Fellows, I. and P. Sing (2012). irr: Various coefficients of interrater reliability and agreement. R Package version 0.84. http://CRAN.R-project.org/package=irr

Appendix 1:

Analysis of Variety Committee Pea CT2

Jean-Louis Laffont, ISTA Statistics Committee

1. Materials and methods

Data are available for 8 laboratories and 24 varieties. Figure 1 summarizes the structure of the data for this CT.

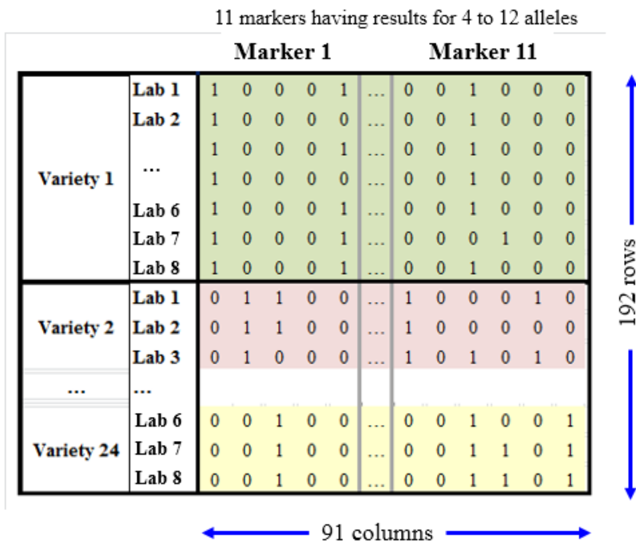

Figure 1: Structure of the data analyzed.

Consider two laboratories and the allele results (0 or 1) for one variety and for the different markers. We elaborate first a coincidence matrix giving the number of 0's and 1's observed in both laboratories and the number of 0's and 1's observed only in one of the two laboratories. Table 1 is an example of such a table. There is a total of 41 alleles and the two laboratories provide same results on 30 + 6 = 36 alleles leading to an overall percent agreement of 36 / 41 = 87.8%. However, this percent agreement is overestimated as agreement between the two laboratories can be due by chance only. That is the reason why many reliability measures taking into account the possible chance agreement have been developed. The most popular one for two laboratories is the Cohen's kappa coefficient (Cohen, 1960). It is computed as:

$$\hat{\kappa}_C = \frac{p_a - p_e}{1 - p_e}$$

in which:

. p_a is the overall percent agreement: $p_a = \frac{n_{00} + n_{11}}{n}$ using the notations in Table 2.

. p_e is the chance agreement probability computed by Cohen (1960) as

$p_e = \frac{n_{0+}}{n} \times \frac{n_{+0}}{n} + \frac{n_{1+}}{n} \times \frac{n_{+1}}{n}$ (other authors have proposed some other ways to compute p_e; see Krippendorff, 2004, for a review of the different ways of computing p_e). In the example, this gives

$p_e = \frac{32}{41} \times \frac{33}{41} + \frac{9}{41} \times \frac{8}{41} = 0.671$ and therefore $\hat{\kappa}_C = \frac{0.878 - 0.671}{1 - 0.671} = 0.629$.

The range of possible values of $\hat{\kappa}_C$ is from -1 to 1. A value of 1 represents perfect agreement, 0 indicates agreement no better than that expected by chance, and a negative value indicates an agreement worse than that expected by chance (Sim and Wright, 2005). Although there is no universally accepted magnitude guidelines on the value of $\hat{\kappa}_C$ for characterizing agreement, we can use the ones established by Landis and Koch (1977) which are summarized in Table 3.

2. Results

Overall percentage agreements (p_a) and Cohen's kappas have been computed for all the possible laboratory pairs and considering as units either the marker alleles or the varieties. The computations have been performed with the R *irr* package (Gamer *et al*., 2012) which includes functions for computing various coefficients of reliability of agreement.

Figures 2 and 3 visualize with dot plots the overall percentages agreements; Figures 4 and 5 visualize Cohen's kappas. Tables 4 and 5 provide the percentages of varieties and the percentage of alleles respectively having a $\hat{\kappa}_C$ value falling in one of the categories defined by Landis and Koch (1977).

The overall agreement percentage for scoring varieties across laboratories is above 90% for all the varieties except one (variety 11) when comparing laboratory 2 with laboratory 5 (p_a = 89%). Most of Cohen's kappa values are above 0.6 for scoring varieties and only few Cohen's kappa values are below 0.6 for scoring alleles.

Given these results, there is enough evidence for validating the method.

		Lab 2		
		0	**1**	**Total**
Lab 1	**0**	30	2	32
	1	3	6	9
	Total	33	8	41

Table 1: Coincidence matrix example.

		Lab 2		
		0	**1**	**Total**
Lab 1	**0**	n_{00}	n_{01}	n_{0+}
	1	n_{10}	n_{11}	n_{1+}
	Total	n_{+0}	n_{+1}	N

Table 2: Abstract coincidence matrix.

$\hat{\kappa}_C$	**Interpretation**
< 0	No agreement
0.0 — 0.20	Slight agreement
0.21 — 0.40	Fair agreement
0.41 — 0.60	Moderate agreement
0.61 — 0.80	Substantial agreement
0.81 — 1.00	Almost perfect agreement

Table 3: Landis and Koch table for the interpretation of $\hat{\kappa}_C$.

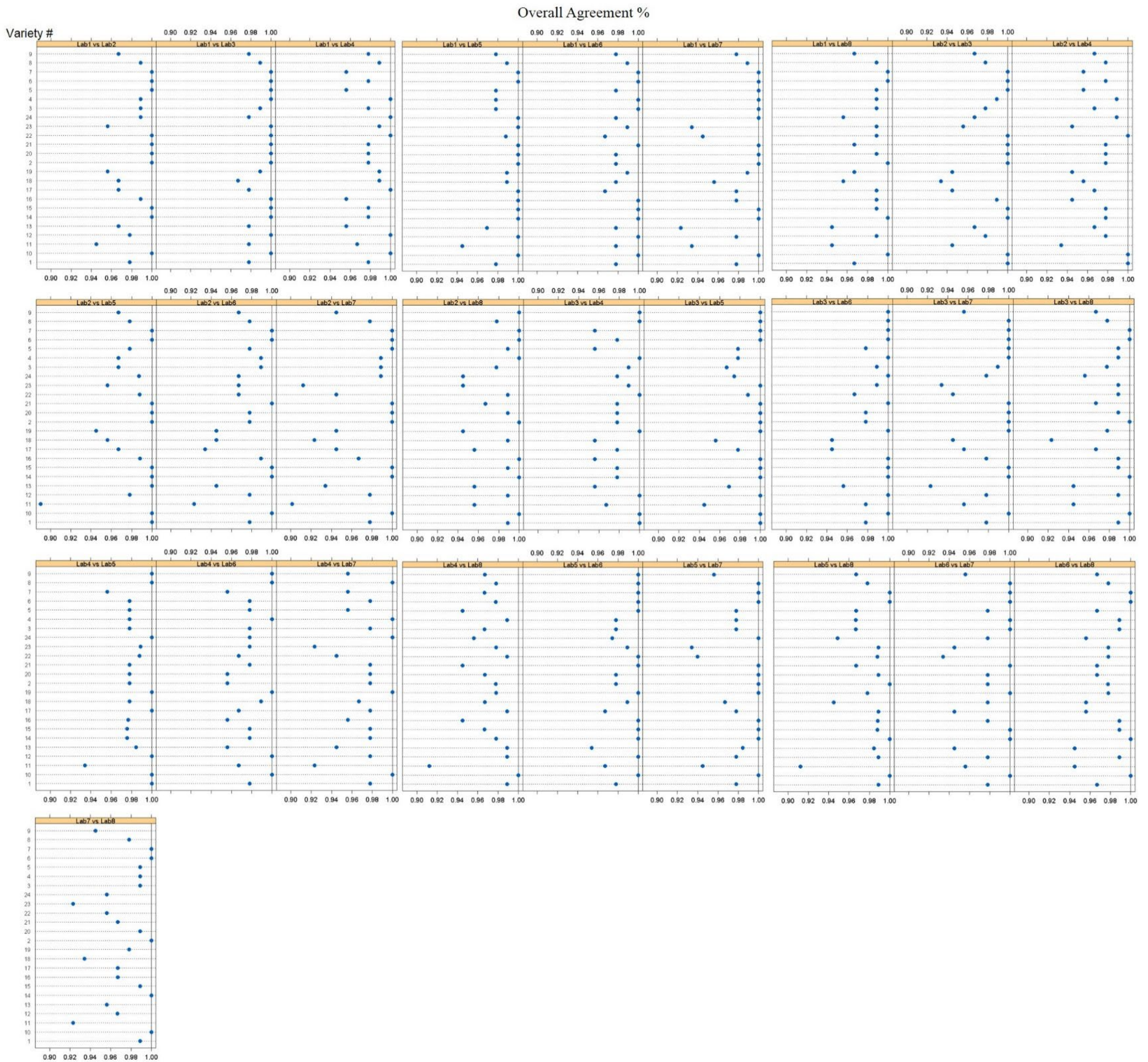

Figure 2: Overall percentage agreements considering allele results agreement for a given variety.

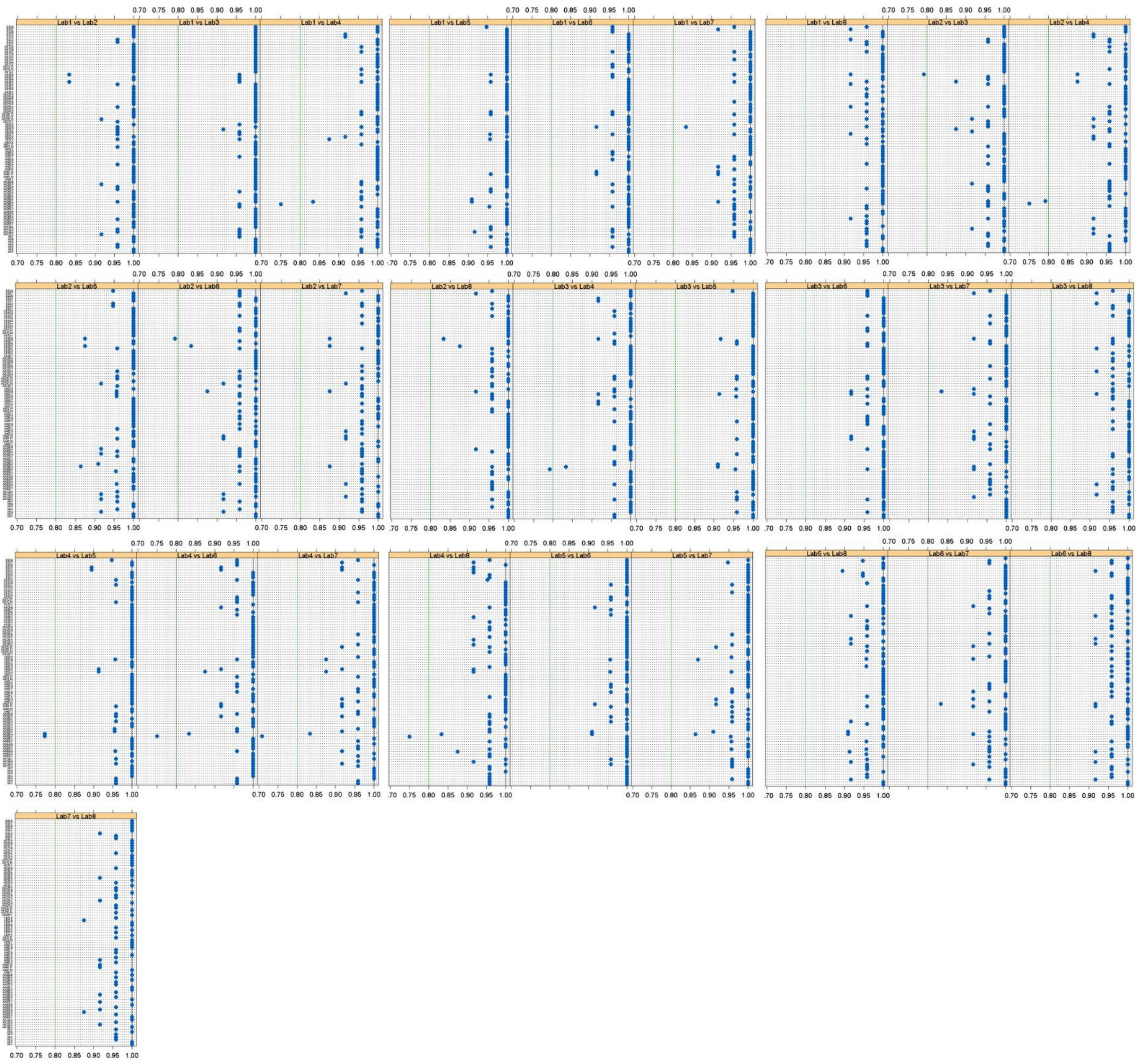

Figure 3: Overall percentage agreements considering allele results agreement across varieties for a given allele.

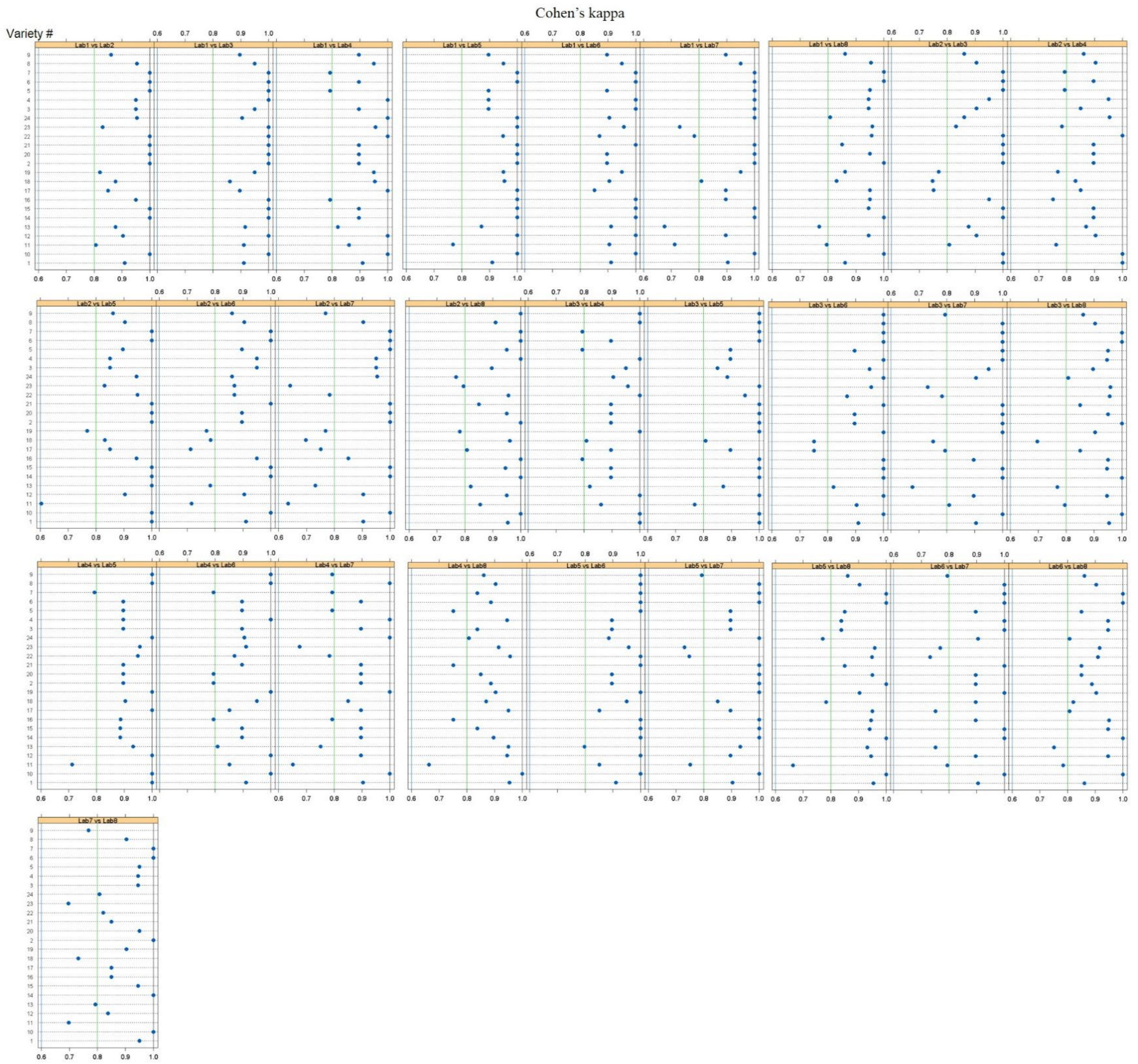

Figure 4: Cohen's kappas considering allele results agreement for a given variety.

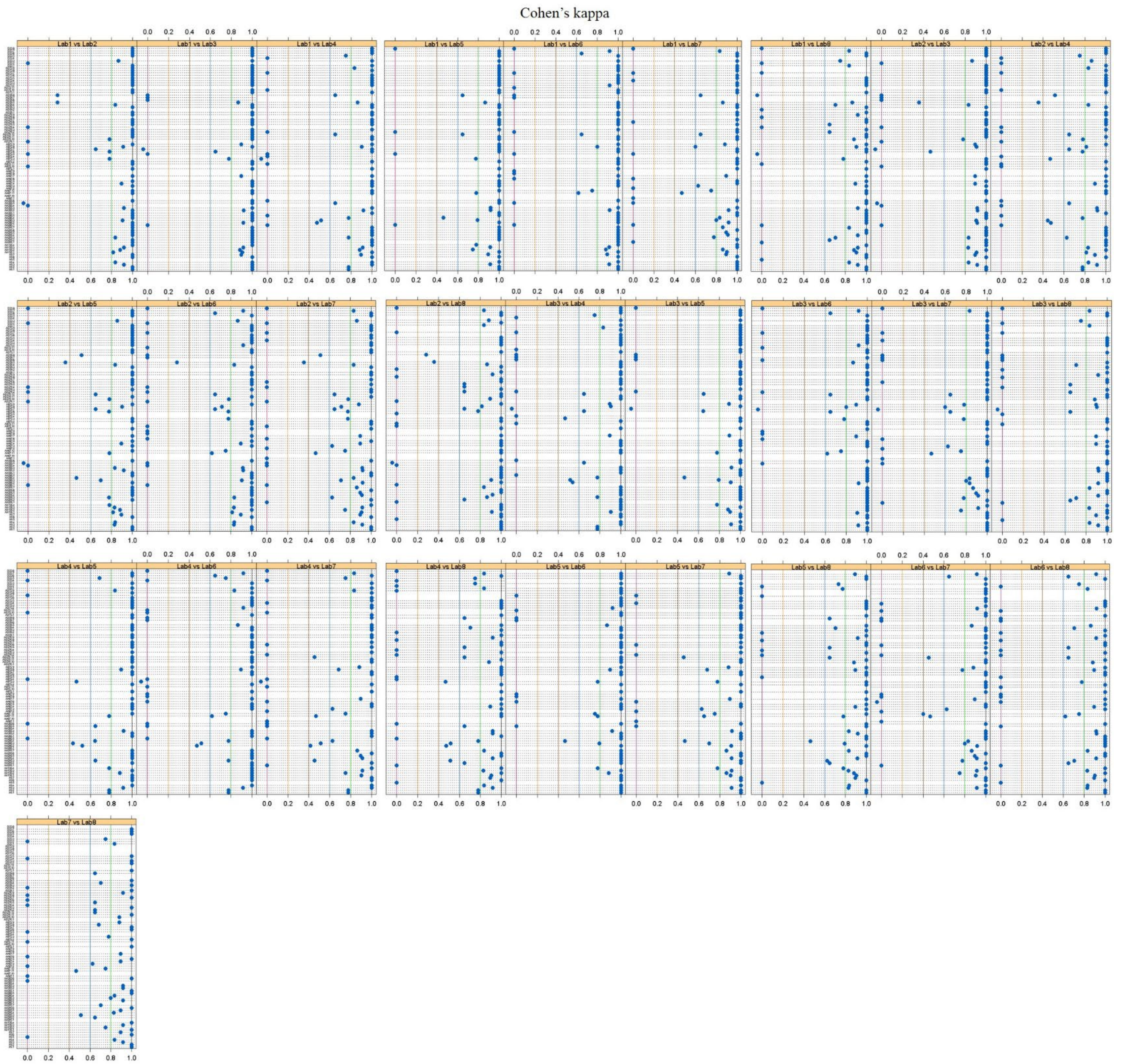

Figure 5: Cohen's kappas considering allele results agreement across varieties for a given allele.

Table 4: For each pair of laboratories, percentage of varieties having a $\hat{\kappa}_C$ value falling in one of the categories defined by Landis and Koch (1977).

% of varieties for which $\hat{\kappa}_C$ is	< 0	0 – 0.2	0.21 – 0.40	0.41 – 0.60	0.61 – 0.80	0.81– 1.00
Lab1 vs Lab2	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab1 vs Lab3	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab1 vs Lab4	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab1 vs Lab5	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab1 vs Lab6	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab1 vs Lab7	0.00%	0.00%	0.00%	0.00%	16.67%	83.33%
Lab1 vs Lab8	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab2 vs Lab3	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab2 vs Lab4	0.00%	0.00%	0.00%	0.00%	25.00%	75.00%
Lab2 vs Lab5	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab2 vs Lab6	0.00%	0.00%	0.00%	0.00%	20.83%	79.17%
Lab2 vs Lab7	0.00%	0.00%	0.00%	0.00%	33.33%	66.67%
Lab2 vs Lab8	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab3 vs Lab4	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab3 vs Lab5	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab3 vs Lab6	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab3 vs Lab7	0.00%	0.00%	0.00%	0.00%	25.00%	75.00%
Lab3 vs Lab8	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab4 vs Lab5	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab4 vs Lab6	0.00%	0.00%	0.00%	0.00%	16.67%	83.33%
Lab4 vs Lab7	0.00%	0.00%	0.00%	0.00%	33.33%	66.67%
Lab4 vs Lab8	0.00%	0.00%	0.00%	0.00%	16.67%	83.33%
Lab5 vs Lab6	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab5 vs Lab7	0.00%	0.00%	0.00%	0.00%	16.67%	83.33%
Lab5 vs Lab8	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab6 vs Lab7	0.00%	0.00%	0.00%	0.00%	25.00%	75.00%
Lab6 vs Lab8	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab7 vs Lab8	0.00%	0.00%	0.00%	0.00%	20.83%	79.17%

Table 5: For each pair of laboratories, percentage of alleles having a $\hat{\kappa}_C$ value falling in one of the categories defined by Landis and Koch (1977).

% of alleles for which $\hat{\kappa}_C$ is	< 0	0 – 0.2	0.21 – 0.40	0.41 – 0.60	0.61 – 0.80	0.81– 1.00
Lab1 vs Lab2	1.30%	7.79%	2.60%	0.00%	5.19%	83.12%
Lab1 vs Lab3	1.35%	6.76%	0.00%	0.00%	2.70%	89.19%
Lab1 vs Lab4	1.35%	13.51%	0.00%	2.70%	10.81%	71.62%
Lab1 vs Lab5	0.00%	5.71%	0.00%	1.43%	10.00%	82.86%
Lab1 vs Lab6	0.00%	15.58%	0.00%	0.00%	6.49%	77.92%
Lab1 vs Lab7	0.00%	14.67%	0.00%	1.33%	9.33%	74.67%
Lab1 vs Lab8	2.74%	13.70%	0.00%	0.00%	9.59%	73.97%
Lab2 vs Lab3	2.56%	11.54%	1.28%	1.28%	1.28%	82.05%
Lab2 vs Lab4	0.00%	17.95%	1.28%	5.13%	12.82%	62.82%
Lab2 vs Lab5	1.37%	9.59%	1.37%	2.74%	10.96%	73.97%
Lab2 vs Lab6	0.00%	19.75%	1.23%	0.00%	11.11%	67.90%
Lab2 vs Lab7	0.00%	18.99%	1.27%	2.53%	13.92%	63.29%
Lab2 vs Lab8	1.32%	15.79%	2.63%	0.00%	7.89%	72.37%
Lab3 vs Lab4	1.35%	14.86%	0.00%	4.05%	10.81%	68.92%
Lab3 vs Lab5	1.43%	7.14%	0.00%	1.43%	7.14%	82.86%
Lab3 vs Lab6	1.33%	12.00%	0.00%	0.00%	9.33%	77.33%
Lab3 vs Lab7	1.33%	16.00%	0.00%	1.33%	13.33%	68.00%
Lab3 vs Lab8	1.35%	17.57%	0.00%	0.00%	9.46%	71.62%
Lab4 vs Lab5	0.00%	10.14%	0.00%	4.35%	11.59%	73.91%
Lab4 vs Lab6	1.33%	18.67%	0.00%	2.67%	10.67%	66.67%
Lab4 vs Lab7	1.35%	20.27%	0.00%	6.76%	10.81%	60.81%
Lab4 vs Lab8	0.00%	19.18%	0.00%	5.48%	15.07%	60.27%
Lab5 vs Lab6	0.00%	11.27%	0.00%	1.41%	7.04%	80.28%
Lab5 vs Lab7	0.00%	11.43%	0.00%	2.86%	10.00%	75.71%
Lab5 vs Lab8	0.00%	11.59%	0.00%	1.45%	15.94%	71.01%
Lab6 vs Lab7	1.35%	14.86%	1.35%	2.70%	8.11%	71.62%
Lab6 vs Lab8	0.00%	21.33%	0.00%	0.00%	13.33%	65.33%
Lab7 vs Lab8	0.00%	18.06%	0.00%	2.78%	19.44%	59.72%

References

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement **20** (1): 37–46.

Gamer, M., Lemon, J., Fellows, I. and P. Sing (2012). irr: Various coefficients of interrater reliability and agreement. R Package version 0.84. http://CRAN.R-project.org/package=irr.

Krippendorff, K. (2004). Reliability in content analysis: some common misconceptions and recommendations. Human Communication Research **30** (3), 411-433.

Landis, J.R. and G.G. Koch (1977). The measurement of observer agreement for categorical data. Biometrics **33** (1): 159–174.

Sim, J. and C.C. Wright (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical Therapy **85** (3): 257–26

Appendix 2

Within laboratory agreement for marker data: proposal

Jean-Louis Laffont

For checking within laboratory agreement (i.e. repeatability) from Variety Committee Comparative Tests (CT), we have alleles from *k* markers on *n* seeds from the same variety. Similar to what is done in Langton *et al*. (2002), the equivalent of repeatability, referred as *accordance*, is defined as the probability that two seeds from the same variety will give both the same allele for a given marker. For marker *i*, it is computed as:

$$p_i = \frac{\sum_{j=1}^{J} \binom{n_j}{2}}{\binom{n}{2}} = \frac{\sum_{j=1}^{J} n_j(n_j - 1)}{n(n-1)}$$

where n_j is the number of seeds with allele *j* (*j* = 1, 2, …, *J*; $n = \sum_{j=1}^{J} n_j$).

The *accordance* for a given variety is then computed as the average of the p_i's over the markers:

$$p = \frac{1}{k}\sum_{i=1}^{k} p_i \ .$$

Table 1 illustrates the computations for five seeds and five markers. When all the alleles are identical, *accordance* is equal to 100% (marker 1) whereas when there are all different, *accordance* is equal to 0% (marker 5). The overall *accordance* for this example is equal to 40%.

Table 1. *Accordance* for the five marker results on five seeds.

Seed number	Marker 1	Marker 2	Marker 3	Marker 4	Marker 5
1	A1	A2	A3	A4	A5
2	A1	A2	A3	A4	B5
3	A1	A2	A3	B4	C5
4	A1	A2	B3	C4	D5
5	A1	B2	C3	D4	E5
n_j	5	4, 1	3, 1, 1	2, 1, 1, 1	1, 1, 1, 1, 1
p_i	1	0.6	0.3	0.1	0

Tables 2 and 3 give the *accordance* for CT 1 on Pea and Oat. For each of the two species, there were 11 marker results for ten individual seeds from eight varieties in two laboratories. The average *accordance* across varieties for the two species is high (above 90%).

Table 2. *Accordance* for Pea.

Variety	Laboratory 1	Laboratory 2
Alfetta	95.8%	95.8%
Cruiser	100.0%	98.2%

Miko	96.8%	96.8%
Mystique	93.5%	93.3%
Nette 2010	100.0%	100.0%
Tamis	90.3%	93.9%
Venture	100.0%	94.5%
Yellowhead	89.1%	94.5%
Mean	95.7%	95.9%

Table 3. *Accordance* for Oat.

Variety	Laboratory 1	Laboratory 2
A	62.0%	91.7%
AC Gwen	91.9%	90.9%
AC Hunter	100.0%	94.5%
B	98.2%	100.0%
C	100.0%	100.0%
Jerry	93.4%	100.0%
Morton	94.9%	86.9%
Souris	85.1%	84.8%
Mean	90.7%	93.6%

Reference

Langton, S.D., Chevennement, R., Nagelkerke, N. and B. Lombard (2002). Analysing collaborative trials for qualitative microbiological methods: accordance and concordance. International Journal of Food Microbiology ***79****, 175–181.*

Appendix 3

Pooled Laboratory Results

SSR name	Allele #	Allele scoring			Variety and laboratory																							
		1	2	3	Alfetta			Cruiser			Miko			Mystique			Nette 2010			Tamis			Venture			Yellowhead		
					1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
A9	1	381		379	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	2	369			0	0	0	0		0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	3	367		365	0	0	0	0		0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	4	365		361?/362	1	1	1	0		0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0
	5	347		344	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
	6	Null			0	0	0	1		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA67	1	385		384	1	1	1	1	1	1	1	1	1	0		0	0	0	0	0	0	1	1	1	1	1	1	1
	2	375	377		0	0	0	0	0	0	0	0	0	1		1	1	1	1	0	0	0	0	0	0	0	0	0
	3	Null			0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	0	0	0	0	0	0	0
AA135	1	368		366	0	0	1	0	0	0	0	0	0	0		0	1	1	1	1	1	1	1	1	1	1	1	1
	2	366		363	1	1	1	1	1	1	0	0	0	1		1	0	0	0	0	0	0	0	0	0	0	0	0
	3	364		362	0	0	0	0	0	0	1	1	1	0		0	0	0	0	0	0	0	0	0	0	0	0	0
AA205	1	241			0	0	0	0	0	0	0	0	0	0		0	0		0	1	1	1	0	0	0	0	0	0
	2	239			0	0	0	0	0	0	0	0	0	0		0	0		0	1	1	1	0	0	0	0	0	0
	3	235			1	1	1	0	0	0	1	1	1	1		1	1		1	0	0	0	1	1	1	0	0	0
	4	233			0	0	0	0	0	0	0	0	0	0	0		0	0		0	0	0	0	0	0	1	1	1
	5	231			0	0	0	1	1	1	0	0	0	0		0	0		0	0	0	0	0	0	0	0	0	0
AA285	1	257/256	256	257	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	2			253	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3	250		251	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
	4	248			1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0
	5	230	232	232	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA355	1	256			0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	248	245	244	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3			246	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	242	239	239	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
	5	190		189	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	6	184			1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
	7	182			0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0
	8	180			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AB72	1	274			0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
	2	270			0	0	0	0	0	0	1	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
	3	268			0	0	0	1	1	1	1	1	1	0	0	0	0		0	0	0	0	1	1	1	1	1	1
	4	266			1	1	1	0	0	0	0	0	0	1	1	1	1		1	0	0	0	0	0	0	0	0	0
	5	264			1	1	1	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
	6	254	256		0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	1	0	0	0	0	0	0
AD59	1	333			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
	2	323			0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0

	3	319		318	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AD73	1		279	272/274	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	272	273	268	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
	3	270			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	245	243	241	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	243	241	239	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0
	6	241		237	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	237	235	233	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
	8	235	233	231	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	9	229	227	225	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
AD270	1	316		317	0	0	1	0	0	0	0	0	0	0		0	0	0	0	0	0	0	1	1	1	0	0	0
	2	314			0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0
	3	306		307	1	1	1	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0
	4	298	300		0	0	0	0	0	0	0	0	0	1		1	0	0	0	0	0	0	0	0	0	0	0	0
	5	286			0	0	0	0	0	0	0	0	0	0		0	1	1	1	0	0	0	0	0	0	0	0	0
	6	282			0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	1	1
	7	278	280		0	0	0	0	0	0	1	1	1	0		0	0	0	0	0	0	0	0	0	0	0	0	0
	8	276			0	0	0	0	0	0	0	0	0	0		0	0	0	0	1	1	1	0	0	1	0	0	0
	9	254	250	252	0	0	0	1	1	1	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0
D23	1	197		196	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
	2	195		192	1	1	1	1	1	1	0	0	0	1	1	1	1		1	0	0	0	1	1	1	0	0	0
	3	193		190	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
	4			188	0	0	0	0	0	1	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
	5	189		186	0	0	0	0	0	0	1	1	1	0	0	0	0		0	0	0	0	0	0	0	0	0	0
	6	177		175	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	1	1	0	0	1	0	0	0
	7	175		173	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	1	1	1

Appendix 4

Analysis of Variety Committee Oat CT2

Jean-Louis Laffont, ISTA Statistics Committee

Materials and methods

Data are available for 7 laboratories and 24 varieties. Figure 1 summarizes the structure of the data for this CT.

9 markers having results for 3 to 10 alleles

		Marker 1						Marker 9					
Variety 1	Lab 1	1	0	0	0	1	...	0	0	1	0	0	0
	Lab 2	1	0	0	0	0	...	0	0	1	0	0	0
	Lab 3	1	0	0	0	1	...	0	0	1	0	0	0
	Lab 4	1	0	0	0	0	...	0	0	1	0	0	0
	Lab 5	1	0	0	0	1	...	0	0	1	0	0	0
	Lab 6	1	0	0	0	1	...	0	0	0	1	0	0
	Lab 7	1	0	0	0	1	...	0	0	1	0	0	0
Variety 2	Lab 1	0	1	1	0	0	...	1	0	0	0	1	0
	Lab 2	0	1	1	0	0	...	1	0	0	0	0	0
	Lab 3	0	1	0	0	0	...	1	0	1	0	1	0
...	...												
Variety 24	Lab 5	0	0	1	0	0	...	0	0	1	0	0	1
	Lab 6	0	0	1	0	0	...	0	0	1	1	0	1
	Lab 7	0	0	1	0	0	...	0	0	1	1	0	1

168 rows

62 columns

Figure 1: Structure of the data analyzed.

Consider two laboratories and the allele results (0 or 1) for one variety and for the different markers. We elaborate first a coincidence matrix giving the number of 0's and 1's observed in both laboratories and the number of 0's and 1's observed only in one of the two laboratories. Table 1 is an example of such a table. There is a total of 41 alleles and the two laboratories provide same results on 30 + 6 = 36 alleles leading to an overall percent agreement of 36 / 41 = 87.8%. However, this percent agreement is overestimated as agreement between the two laboratories can be due by chance only. That is the reason why many reliability measures taking into account the possible chance agreement have been developed. The most popular one for two laboratories is the Cohen's kappa coefficient (Cohen, 1960). It is computed as:

$$\hat{\kappa}_C = \frac{p_a - p_e}{1 - p_e}$$

in which:

. p_a is the overall percent agreement: $p_a = \frac{n_{00} + n_{11}}{n}$ using the notations in Table 2.

. p_e is the chance agreement probability computed by Cohen (1960) as

$p_e = \frac{n_{0+}}{n} \times \frac{n_{+0}}{n} + \frac{n_{1+}}{n} \times \frac{n_{+1}}{n}$ (other authors have proposed some other ways to compute p_e; see Krippendorff, 2004, for a review of the different ways of computing p_e). In the example, this gives $p_e = \frac{32}{41} \times \frac{33}{41} + \frac{9}{41} \times \frac{8}{41} = 0.671$ and therefore $\hat{\kappa}_C = \frac{0.878 - 0.671}{1 - 0.671} = 0.629$.

The range of possible values of $\hat{\kappa}_C$ is from -1 to 1. A value of 1 represents perfect agreement, 0 indicates agreement no better than that expected by chance, and a negative value indicates an agreement worse than that expected by chance (Sim and Wright, 2005). Although there is no universally accepted magnitude guidelines on the value of $\hat{\kappa}_C$ for characterizing agreement, we can use the ones established by Landis and Koch (1977) which are summarized in Table 3.

Results

Overall percentage agreements (p_a) and Cohen's kappas have been computed for all the possible laboratory pairs and considering as units either the marker alleles or the varieties. The computations have been performed with the R *irr* package (Gamer *et al*., 2012) which includes functions for computing various coefficients of reliability of agreement.

Figures 2 and 3 visualize with dot plots the overall percentages agreements; Figures 4 and 5 visualize Cohen's kappas. Tables 4 and 5 provide the percentages of varieties and the percentage of alleles respectively having a $\hat{\kappa}_C$ value falling in one of the categories defined by Landis and Koch (1977).

The overall agreement percentage for scoring varieties across laboratories is above 90% for all the varieties except one (variety 16) when comparing laboratory 4 with laboratory 7 (p_a = 88.7%). Most of Cohen's kappa values are above 0.8 for scoring varieties and only few Cohen's kappa values are below 0.6 for scoring alleles.

Given these results, there is enough evidence for validating the method.

		Lab 2		
		0	1	Total
Lab 1	0	30	2	32
	1	3	6	9
	Total	33	8	41

Table 1: Coincidence matrix example.

		Lab 2		
		0	1	Total
Lab 1	0	n_{00}	n_{01}	n_{0+}
	1	n_{10}	n_{11}	n_{1+}
	Total	n_{+0}	n_{+1}	N

Table 2: Abstract coincidence matrix.

$\hat{\kappa}_C$	**Interpretation**
< 0	No agreement
0.0 — 0.20	Slight agreement
0.21 — 0.40	Fair agreement
0.41 — 0.60	Moderate agreement
0.61 — 0.80	Substantial agreement
0.81 — 1.00	Almost perfect agreement

Table 3: Landis and Koch table for the interpretation of $\hat{\kappa}_C$.

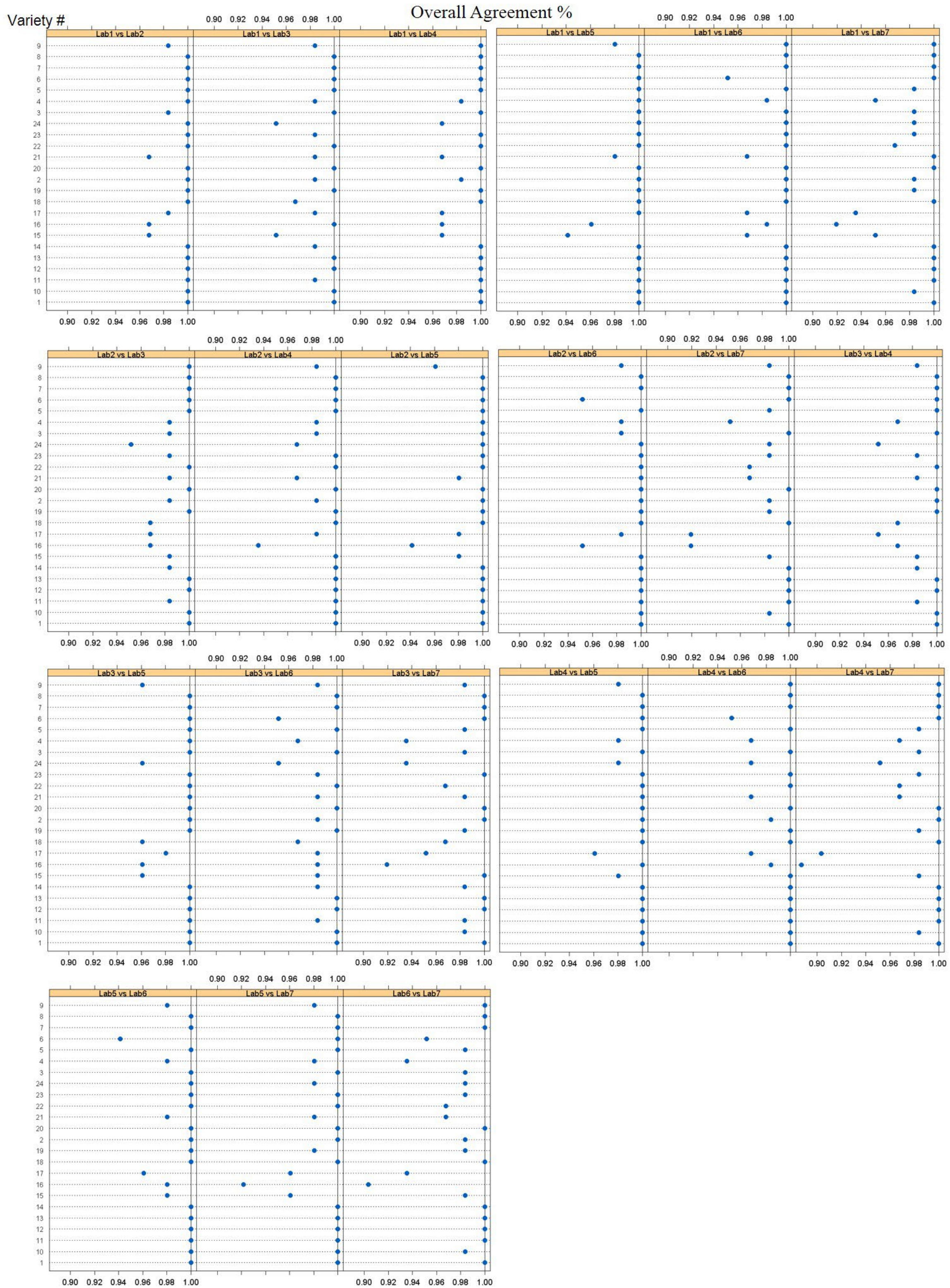

Figure 2: Overall percentage agreements considering allele results agreement for a given variety.

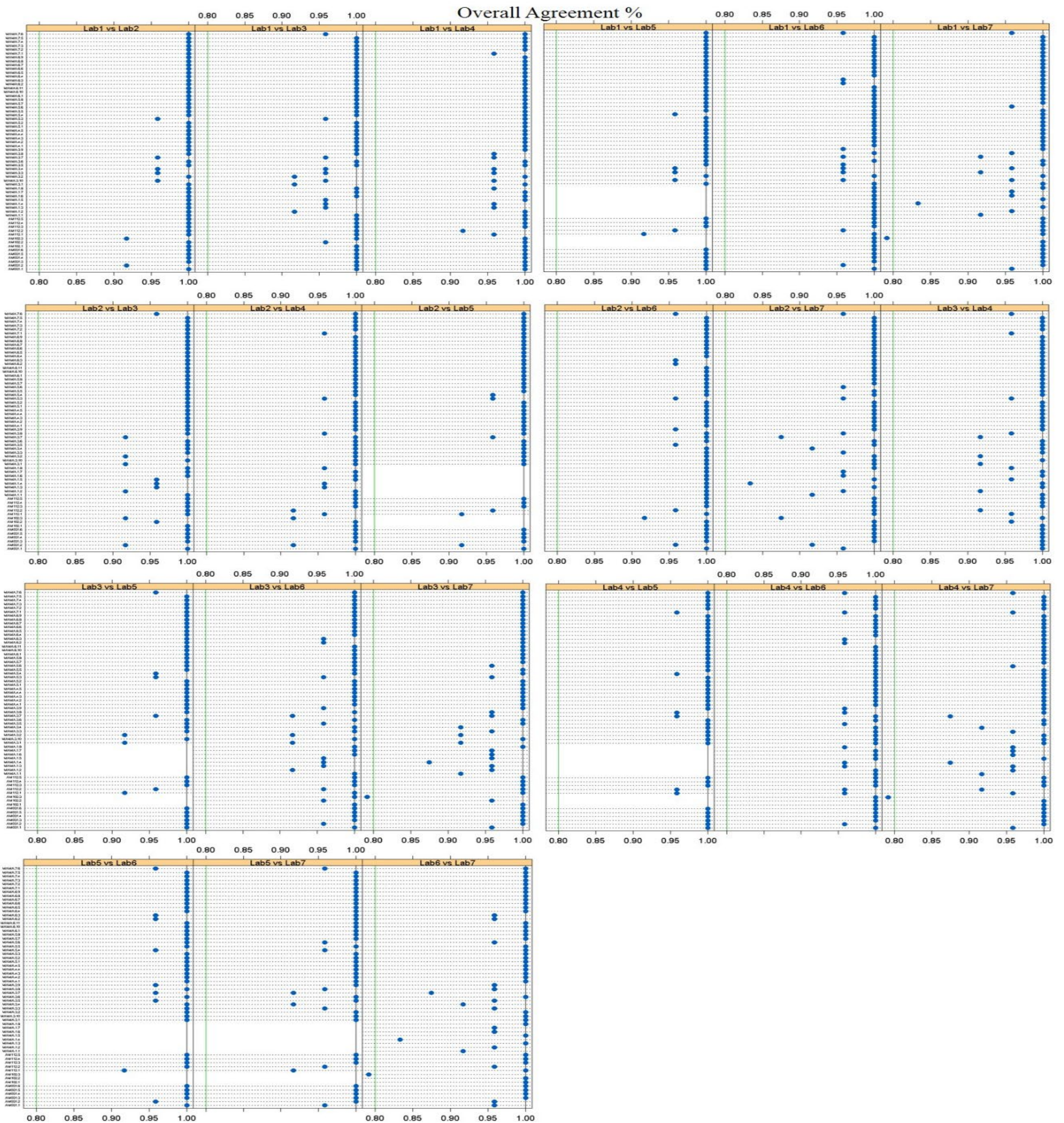

Figure 3: Overall percentage agreements considering allele results agreementacross varieties for a given allele.

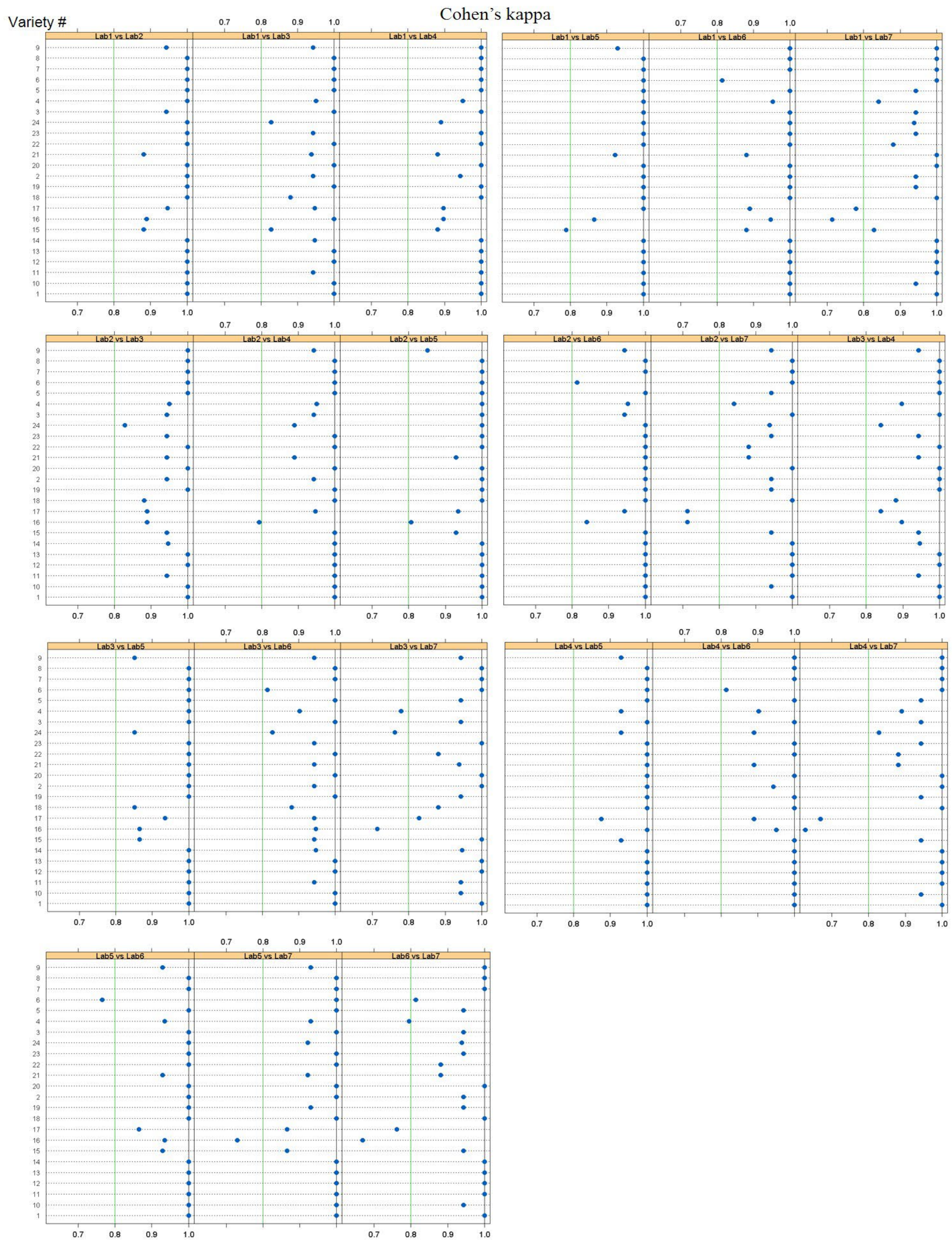

Figure 4: Cohen's kappas considering allele results agreement for a given variety.

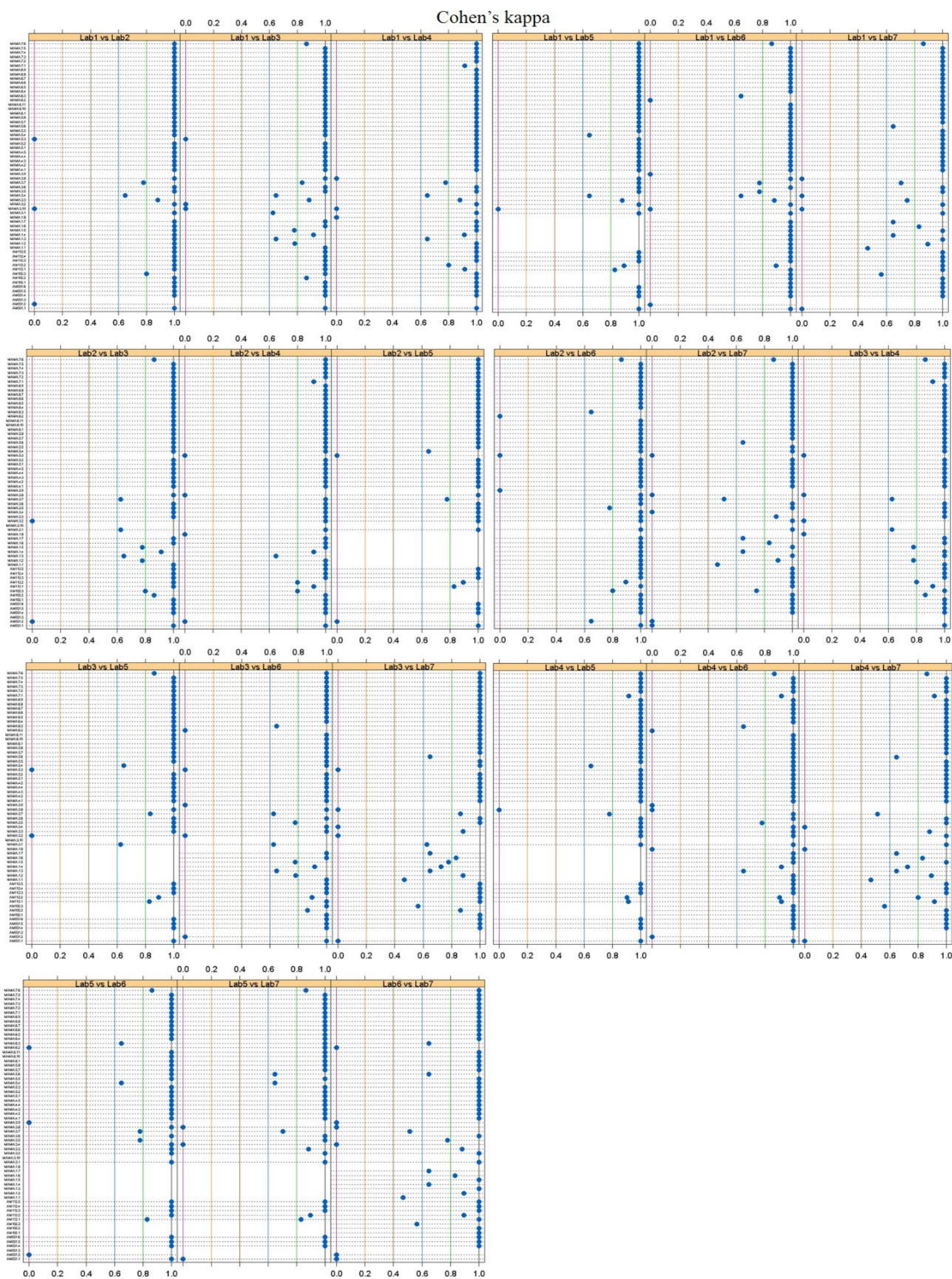

Figure 5: Cohen's kappas considering allele results agreement across varieties for a given allele.

Table 4: For each pair of laboratories, percentage of varieties having a $\hat{\kappa}_C$ value falling in one of the categories defined by Landis and Koch (1977).

% of varieties for which $\hat{\kappa}_C$ is	< 0	0 – 0.2	0.21 – 0.40	0.41 – 0.60	0.61 – 0.80	0.81– 1.00
Lab1 vs Lab2	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab1 vs Lab3	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab1 vs Lab4	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%

Lab1 vs Lab5	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab1 vs Lab6	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab1 vs Lab7	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab2 vs Lab3	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab2 vs Lab4	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab2 vs Lab5	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab2 vs Lab6	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab2 vs Lab7	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab3 vs Lab4	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab3 vs Lab5	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab3 vs Lab6	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab3 vs Lab7	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%
Lab4 vs Lab5	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab4 vs Lab6	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
Lab4 vs Lab7	0.00%	0.00%	0.00%	0.00%	8.33%	91.67%
Lab5 vs Lab6	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab5 vs Lab7	0.00%	0.00%	0.00%	0.00%	4.17%	95.83%
Lab6 vs Lab7	0.00%	0.00%	0.00%	0.00%	12.50%	87.50%

Table 5: For each pair of laboratories, percentage of alleles having a $\hat{\kappa}_C$ value falling in one of the categories defined by Landis and Koch (1977).

% of alleles for which $\hat{\kappa}_C$ is	< 0	0 – 0.2	0.21 – 0.40	0.41 – 0.60	0.61 – 0.80	0.81– 1.00
Lab1 vs Lab2	0.00%	5.08%	0.00%	0.00%	5.08%	89.83%
Lab1 vs Lab3	0.00%	5.17%	0.00%	0.00%	8.62%	86.21%
Lab1 vs Lab4	0.00%	5.08%	0.00%	0.00%	6.78%	88.14%
Lab1 vs Lab5	0.00%	2.08%	0.00%	0.00%	4.17%	93.75%
Lab1 vs Lab6	0.00%	6.67%	0.00%	0.00%	6.67%	86.67%
Lab1 vs Lab7	0.00%	6.90%	0.00%	3.45%	8.62%	81.03%
Lab2 vs Lab3	0.00%	3.51%	0.00%	0.00%	10.53%	85.96%
Lab2 vs Lab4	0.00%	6.78%	0.00%	0.00%	5.08%	88.14%
Lab2 vs Lab5	0.00%	4.17%	0.00%	0.00%	4.17%	91.67%
Lab2 vs Lab6	0.00%	5.08%	0.00%	0.00%	6.78%	88.14%
Lab2 vs Lab7	0.00%	8.62%	0.00%	3.45%	6.90%	81.03%
Lab3 vs Lab4	0.00%	6.90%	0.00%	0.00%	8.62%	84.48%
Lab3 vs Lab5	0.00%	4.26%	0.00%	0.00%	4.26%	91.49%
Lab3 vs Lab6	0.00%	8.47%	0.00%	0.00%	11.86%	79.66%
Lab3 vs Lab7	0.00%	8.77%	0.00%	3.51%	10.53%	77.19%
Lab4 vs Lab5	0.00%	2.13%	0.00%	0.00%	4.26%	93.62%
Lab4 vs Lab6	0.00%	8.33%	0.00%	0.00%	5.00%	86.67%
Lab4 vs Lab7	0.00%	5.26%	0.00%	5.26%	8.77%	80.70%
Lab5 vs Lab6	0.00%	6.12%	0.00%	0.00%	8.16%	85.71%
Lab5 vs Lab7	0.00%	6.38%	0.00%	0.00%	6.38%	87.23%
Lab6 vs Lab7	0.00%	10.17%	0.00%	5.08%	8.47%	76.27%

References

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20** (1): 37–46.

Gamer, M., Lemon, J., Fellows, I. and P. Sing (2012). irr: Various coefficients of interrater reliability and agreement. R Package version 0.84. http://CRAN.R-project.org/package=irr.

Krippendorff, K. (2004). Reliability in content analysis: some common misconceptions and recommendations. *Human Communication Research* **30** (3), 411-433.

Landis, J.R. and G.G. Koch (1977). The measurement of observer agreement for categorical data. *Biometrics* **33** (1): 159–174.

Sim, J. and C.C. Wright (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy* **85** (3): 257–268.

Appendix 5

Within laboratory agreement for marker data: proposal

Jean-Louis Laffont

For checking within laboratory agreement (i.e. repeatability) from Variety Committee Comparative Tests (CT), we have alleles from *k* markers on *n* seeds from the same variety. Similar to what is done in Langton *et al*. (2002), the equivalent of repeatability, referred as *accordance*, is defined as the probability that two seeds from the same variety will give the same allele for a given marker. For marker *i*, it is computed as:

$$p_i = \frac{\sum_{j=1}^{J} \binom{n_j}{2}}{\binom{n}{2}} = \frac{\sum_{j=1}^{J} n_j(n_j - 1)}{n(n-1)}$$

where n_j is the number of seeds with allele *j* (*j* = 1, 2, … , *J* ; $n = \sum_{j=1}^{J} n_j$).

The *accordance* for a given variety is then computed as the average of the p_i's over the markers:

$$p = \frac{1}{k}\sum_{i=1}^{k} p_i \ .$$

Table 1 illustrates the computations for five seeds and five markers. When all the alleles are identical, *accordance* is equal to 100% (marker 1) whereas when there are all different, *accordance* is equal to 0% (marker 5). The overall *accordance* for this example is equal to 40%.

Table 1. *Accordance* for the five marker results on five seeds.

Seed number	Marker 1	Marker 2	Marker 3	Marker 4	Marker 5
1	A1	A2	A3	A4	A5
2	A1	A2	A3	A4	B5
3	A1	A2	A3	B4	C5
4	A1	A2	B3	C4	D5
5	A1	B2	C3	D4	E5
n_j	5	4, 1	3, 1, 1	2, 1, 1, 1	1, 1, 1, 1, 1
p_i	1	0.6	0.3	0.1	0

Tables 2 and 3 give the *accordance* for CT 1 on Pea and Oat. For each of the two species, there were 11 marker results for ten individual seeds from eight varieties in two laboratories. The average *accordance* across varieties for the two species is high (above 90%).

Table 2. *Accordance* for Pea.

Variety	Laboratory 1	Laboratory 2
Alfetta	95.8%	95.8%
Cruiser	100.0%	98.2%

Miko	96.8%	96.8%
Mystique	93.5%	93.3%
Nette 2010	100.0%	100.0%
Tamis	90.3%	93.9%
Venture	100.0%	94.5%
Yellowhead	89.1%	94.5%
Mean	95.7%	95.9%

Table 3. *Accordance* for Oat.

Variety	Laboratory 1	Laboratory 2
A	62.0%	91.7%
AC Gwen	91.9%	90.9%
AC Hunter	100.0%	94.5%
B	98.2%	100.0%
C	100.0%	100.0%
Jerry	93.4%	100.0%
Morton	94.9%	86.9%
Souris	85.1%	84.8%
Mean	90.7%	93.6%

Reference

Langton, S.D., Chevennement, R., Nagelkerke, N. and B. Lombard (2002). Analysing collaborative trials for qualitative microbiological methods: accordance and concordance. International Journal of Food Microbiology **79**, 175–181.

Appendix 6

Pooled Laboratory Results

SSR	Allele	Allele scoring			Variety a laboratory																							
		1	2	3	A			B			C			Jerry			Morton			Souris			AC Gwen			AC		
					1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
AM031	1	14		158	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
	2	14			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
	3		169	184	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	18		198	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
AM102	1	20		219/2	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	2	21		232	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	3	22		234	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
AM112	1	23		248/2	1	1	1	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	
	2	23		252	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	3	25		272	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	
MAMA	1	19		209	1	1		0	0	0	0	0	0	0	0	0	0	0		1	1	0	0	0		0	0	
	2	20			1	1		1	1	0	0	0	0	0	0	0	0	0		0	0	0	0	0		1	1	
	3	20		217	0	0		0	0	0	1	1	0	1	1	1	0	0		0	0	0	0	0		0	0	
	4	20		222	0	0		0	0	0	0	0	0	0	0	0	0	0		0	0	0	1	1		0	0	
	5	Nu			0	0		0	0	0	0	0	0	0	0	0	1	1		0	0	0	0	0		0	0	
	6			226	0	0		0	0	1	0	0	1	0	0	0	0	0		0	0	1	0	0		0	0	
MAMA	1	35	354	372/3	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	
	2	36	362/3	381/3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	1	
	3	37		391/3	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	
	4	38		400	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	
	5	38		404	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	6	39		407/4	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	
	7	39		412	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	
	8	39		415/4	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	9		Null		0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
MAMA	1		344		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	2	34		360/3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
MAMA	1	16	164		0	0	0	0	0		0	0		0	0	0	0	0	0	1	1		0	0	0	0	0	
	2	21			0	0	0	1	1		1	1		0	0	0	0	0	0	1	1		0	0	0	0	0	

	3	21		230	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0
	4	24			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	5	25		266/2	1	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
	6	25		270/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAMA	1	16			0	0		0	0	0	0	0	0		0	0		0	0	0	0		1	1
	2	17			0	0		0	0	1	1	0	0		0	0		0	0	0	0		0	0
	3	21			0	0		1	1	0	0	0	0		1	1		0	0	1	1		0	0
	4	21			1	1		0	0	0	0	0	0		0	0		1	1	1	1		0	0
	5	21			0	0		0	0	0	0	1	1		0	0		0	0	0	0		0	0
	6	21			0	0		0	0	0	0	0	0		0	0		0	0	0	0		0	0
MAMA	1	32			1	1		0	0	0	0	0	0		0	0		0	0	0	0		1	1
	2	34	345		0	0		1	1	0	0	0	0		0	0		0	0	0	0		0	0
	3	35			0	0		0	0	0	0	1	1		1	1		1	1	1	1		0	0
	4	35			0	0		0	0	0	0	0	0		0	0		0	0	0	0		0	0
	5	35	359		0	0		0	0	1	1	0	0		0	0		0	0	1	1		0	0