INTERNATIONAL SEED TESTING ASSOCIATION (ISTA)
Secretariat, Zürichstrasse 50, 8303 Bassersdorf, CH-Switzerland
Phone: +41-44-838 60 00, Fax: +41-44-838 60 01, Email: ista.office@ista.ch,
http://www.seedtest.org

Interlaboratory PT for detection of *Ustilago nuda* on *Hordeum vulgare*

PT 15 SH U. nuda

ISTA Seed Health Committee

Results of ISTA Seed Health Proficiency Test Round PT15 SH *Ustilago nuda* on *Hordeum vulgare*

Valerie Cockerell[1] and Adrian Roberts[2]

[1]*Ustilago nuda* Proficiency Test Leader, Official Seed Testing Station for Scotland, Science and Advice for Scottish Agriculture (SASA), Edinburgh, EH12 9FJ and [2] Biomathematics & Statistics Scotland, Edinburgh, EH9 3JZ

Introduction

This is the third proficiency test for *Ustilag nuda* on *Hordeum vulgare,* ISTA Method 7-013. Since the last proficiency test in 2008, a new method was introduced in 2012, "*Method 7-013b Detection of Ustilago nuda in Hordeum vulgare (barley) seed by dehulling and embryo extraction*" and the existing method re-numbered to 7-013a. Method 7-013b is considered equivalent to Method 7-013a offering an alternative embryo extraction method for laboratories that do not have a fume hood or access to a plentiful supply of warm water. The aim of the Proficiency Test Round PT15-SH *Ustilago nuda* was to check the ability of laboratories to identify the presence of the fungus *Ustilago nuda* in barley embryos and quantify the number of infected embryos. To help reduce the high variability seen between laboratories in previous proficiency tests examples of infected embryos were provided to all laboratories as a positive control. New Guidelines[1] developed by the ISTA Seed Health and Statistic Committees to help Proficiency Test Leaders co-ordinate and analyse seed health proficiency tests were used where appropriate in the reporting of this proficiency test.

Proficiency Test Organisation

Experimental design
Three seed lots were chosen from commercial samples tested at the Official seed Testing Station for Scotland,SASA: a healthy (nil infection), medium infection (between 0.5% and <1.0%) and a high infection (>1.0%). Infection levels were chosen to best represent seed lots that laboratories are likely to test and at a suffiently high level to determine a laboratory's competence. One sub-sample of 120g from each seed lot was sent to each participating laboratory. Using only one sub-sample deviates from the Guidelines. The Guidelines suggest that a minimum of three be used. However due to the high probability of infection being found in one sub-sample of Lot 2 and Lot 3 the organisers decided to go ahead with one sub-sample, reducing the workload for participating laboratories.

Sample preparation
A total of 100 subsamples of 120g were prepared according to the 'ISTA Proficiency Test Sample Preparation Instruction'. All subsamples were randomly numbered. Twenty-eight laboratories participated, 15 accredited and 13 volunteer.

Testing
Each participating laboratory received three coded subsamples (one subsample from each lot) for testing. Laboratories accredited for *Ustilago nuda* detection were required to use the method within their scope of accreditation. If a laboratory included both methods within their scope then a laboratory received two sets of samples one for each method. Laboratories taking part on a voluntary basis could choose their

[1] Guidelines for organising and analysing results of Proficiency Tests (PT) draft 21-1-13V3.0.

preferred method. Participants were asked to record the number of infected embryos found in an examination of 1000 embryos.

Homogeneity Conformance

The organising laboratory tested 10 randomly chosen subsamples from each seed lot for determination of heterogeniety using Method 7-013a. Homogeniety was checked using the Excel tool devised by ISTA Statistics Committee available at http://seedtest.org/en/tool-box-_content---1--1410.html. Both seed lots 2 and 3 met homogeniety check requirements and could be considered suitable for use in the proficiency test, Tables 1 and 2. No infected embryos were found in the 10 Healthy Lot 1 subsamples, confirming Lot 1 was a suitable Healthy Lot.

Table 1 Homogeniety results for Lot 2

homogeneity test - seed health											
Sample size	1000										
Samples	1	2	3	4	5	6	7	8	9	10	
infected seed nb.	5	7	6	6	10	11	7	10	10	7	Average
healthy seed percentage	99.5	99.3	99.4	99.4	99	98.9	99.3	99	99	99.3	99.21
Average					99.21						
Tolerance					1.41						
H value											
Homogeneity check					OK						

Table 2 Homogeniety results for Lot 3

homogeneity test - seed health											
Sample size	1000										
Samples	1	2	3	4	5	6	7	8	9	10	
infected seed nb.	9	15	9	9	13	11	12	11	11	11	Average
healthy seed percentage	99.1	98.5	99.1	99.1	98.7	98.9	98.8	98.9	98.9	98.9	98.89
Average					98.89						
Tolerance					1.41						
H value											
Homogeneity check					OK						

Stability Tests

Three subsamples were tested from each lot at the end of the Proficiency Test round to determine stability of infection. Stability tests results are shown in Table 3. Results were checked using the Homogeneity tool described above by adding the three new test results to the original ten . All 13 results for each of Lot 2 and Lot 3 were accepted as homogeneous for the purposes of the Proficiency Test.

Table 3. Stability test results

Lot 1 Healthy	Lot 2 Medium	Lot 3 High
0	14	18
0	8	13
0	10	15

Laboratory PT Results

Results were received from 28 laboratories. One laboratory was accredited for both methods and tested two sets of samples. This laboratory has been assigned two random laboratory numbers in this report.

Four laboratories reported infected embryos in the healthy seed lot, Lab 11(578), Lab 24 (3), Lab 28 (1) and Lab 34 (49). According to the Guidelines these are regarded as false positives as all homogeniety and stabillity results for the Healthy Lot 1 showed nil infection.

All laboratory results for Lots 2 and 3 are shown in Figure 1. Laboratory 11 had exceptionally high results for all seed lots suggesting a fundamental problem with the testing. The mean infection and number of false positives for accredited and volunteer laboratories, together with the homogeniety subsample results are given in Table 3. Results from Lab 11 are excluded from Table 3.

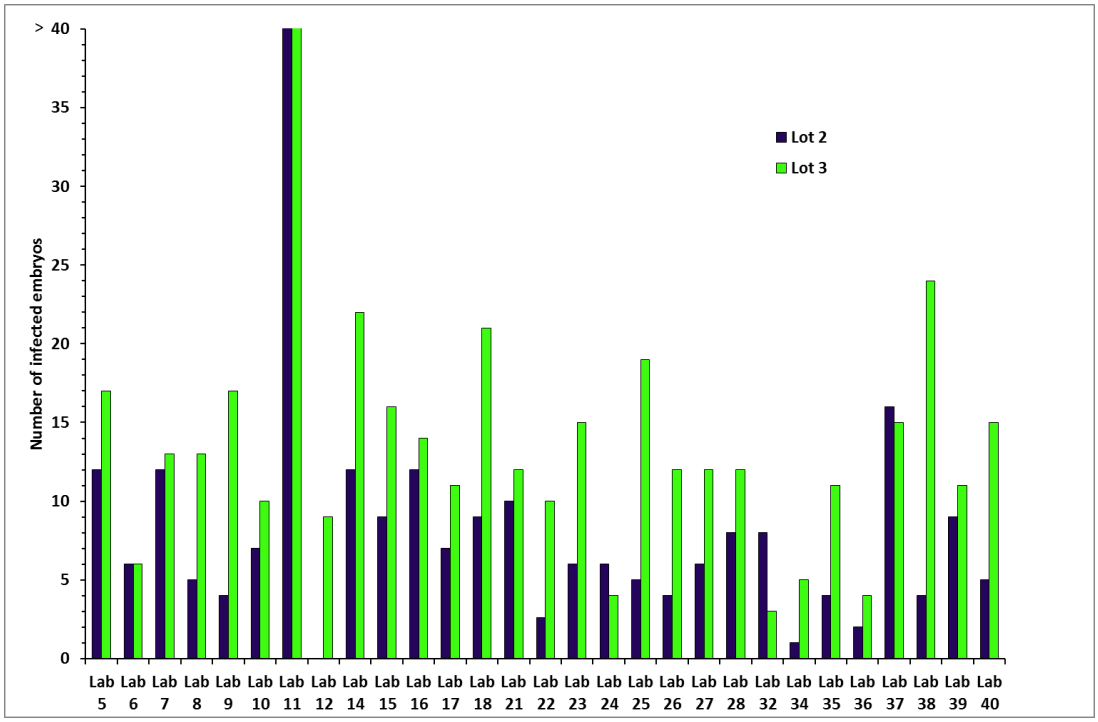


Figure 1 Number of infected embryos found in 1000 embryos examined for Lots 2 and 3 (all laboratories).

Table 3. False positives (Lot 1 healthy), mean proportion infected and range of infection (Lot 2 and Lot 3)*

Laboratory grouping	# of sub-samples with false positives (range)	Mean proportion of embryos (and range of infection) (%)	
	Lot 1 Healthy	Lot 2 Medium	Lot 3 High
Accredited laboratories	1 (3)	0.71 (0.4-1.2)	1.44 (0.4-2.4)
Volunteer laboratories	2 (1-49)	0.65 (0.1-1.6)	1.05 (0.3-2.1)
Homogeneity samples tested by organising laboratory	0 (0)	0.79(0.5-1.1)	1.11(0.9-1.5)

*excludes data from Laboratory 11

As per the 'Guidelines' Hampels test was used to determine any outliers for Lots 2 and 3 after removal of two laboratories (Lab 22 and Lab 27) who had not followed PT instructions and Lab 11. On examination of the data it was felt that there was a problem when counts were low. In particular for lot 2, where a lab that scores for infected embryos zero is not picked up either by the Hampel method or by the Z-score. Given results for Lot 2 it seems likely for a zero count to be an outlier. The fact that it is not, is likely due to low counts overall for this lot and the associated lack of symmetry of the likely sampling distribution (perhaps it is better to use the binomial distribution or a modification of the binomial allowing for greater variability). This should not be an issue for Lot 3 where counts of infected seed were higher.

It was decided to calculate Z-scores for lot 3 using the Excel spreadsheet 'Seed Health Proficiency Test Rating' from the ISTA website (http://seedtest.org/en/tool-box- content---1--1410.html). The mean from Accredited laboratories results were used to determine Z-scores for all laboratories Figure 2. Z-scores larger than 2 or less than -2 are considered to have a possible problem with either over estimating the number of infected embryos (L11) or underestimating the number of infected embryos (L32, L24 and L36) respectively. Overall laboratories were more likely to underestimate the number of infected embryos than overestimate.

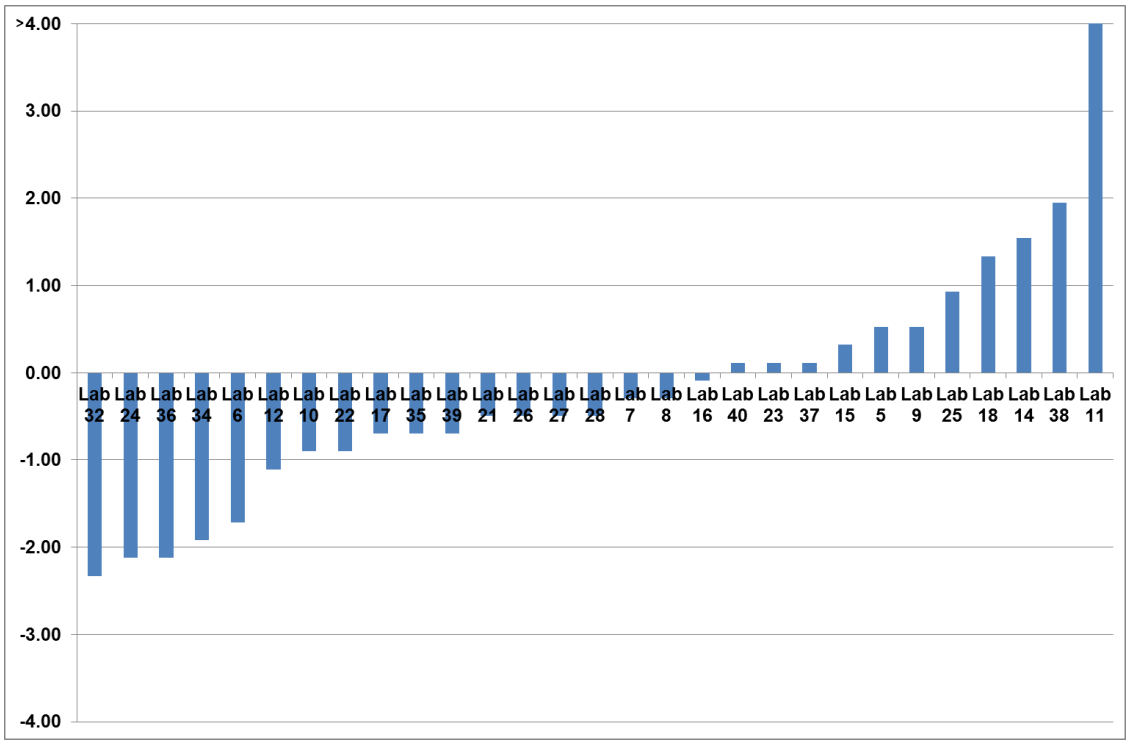


Figure 2. Z-scores for seed Lot 3 all Laboratories

For Lot 2 we decided to assume that the data is binomial with potential outliers. We estimated the proportion for the binomial distribution by the mean of the individual proportions. Using this proportion we looked at how extreme each lab was compared to the expected distribution. This resulted in p-values; the p-value thresholds could be used to categorise laboratory results as good or bad. Figure 3 shows a possible categorisation based on limits of acceptable binomial p-values dependent on limits given by quantile of a normal distribution. In this case a critical value would be <0.01, possible problem ≥0.01<0.1, acceptable results ≥0.1. This showed three laboratories were below the critical value suggesting they had either over-estimated or under-

estimated the number of embryos in their subsample for Lot 2. A further seven laboratories were categorized as having a potential problem. The remaining laboratories had satisfactory results.

Table 4 Using Binomial p-values to categorize Lot 2 laboratory results (number of infected embryos recorded)

Below critical value Over or under estimating number of infected embryos	Possible problem over/under estimating number of infected embryos	Acceptable Results
Labs <0.01	Labs ≥0.01<0.1	Lab ≥0.1
11 (605)	5 (12)	6 (6)
12(0)	7 (12)	8 (5)
37 (16)	14 (12)	9 (4)
	16 (12)	10 (7)
	22 (3)	15 (9)
	34 (1)	17 (7)
	36 (2)	18 (9)
		21 (10)
		23 (6)
		24 (6)
		25 (5)
		26 (4)
		27 (6)
		28 (8)
		32 (8)
		35 (4)
		38 (4)
		39 (9)
		40 (5)
Accredited laboratories mean	7.12	

.

Rating Laboratories

The ISTA Seed Health Committee are at present evaluating methods that could be used to rate laboratories. For now ratings are provided for information and have no consequences on the accreditation of the laboratories.

Ratings have been calculated for Lot 1 and Lot 3 as per the Guidelines using the absolute Z-scores for Lot 3 and for the Healthy Lot 1 as follows:

Health Lot
A: limit is 0 contaminated seeds.
B: accept 1 false positive on total tested and enter the z-score value of 1 false positive as the limit.
C: accept 2 false positive on total tested and enter the z-score value of 2 false positive as the limit.

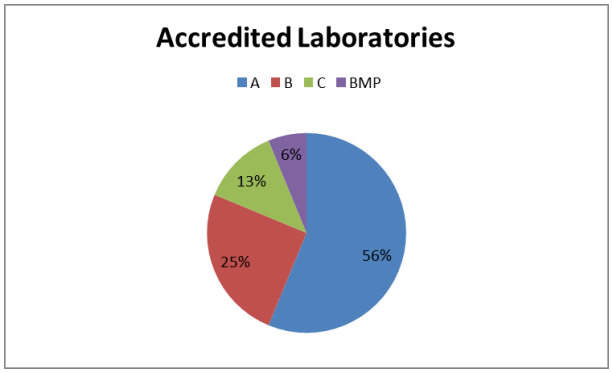
For Contaminated samples –Lot 3
Limits of acceptable Z scores for A, B, C and BMP are dependent on limits given by quantile of a normal distribution:

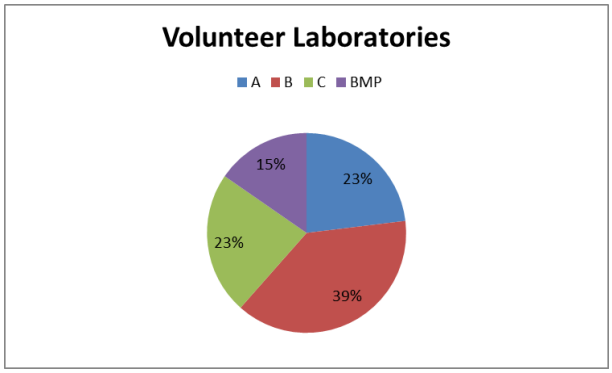
- Limit for A: 0.67 which corresponds to 0.75 quantile of a normal distribution
- Limit for B: mean between A and C
- Limit for C: 2.33 which corresponds to 0.99 quantile of a normal distribution

It will give the values already filled on the sheet: (in absolute values of Z scores)

- A: < 0.67
- B: >0.67 and <1.5
- C: > 1.5 and <2.33
- BMP: > 2.33

Laboratory ratings achieved are summarised in Figure 3. All laboratories results were included. Accredited laboratories received a higher proportion of A ratings and a lower proportion with BMP. Overall 73% of laboratories gave acceptable results, 17% with a possible problem and 10% below minimum required performance. BMP performances were mostly related to false positive results.





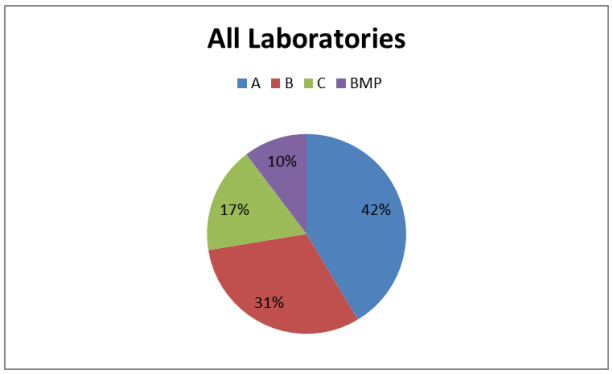


Figure 3 Proportion of Ratings for different laboratory groups. Number of Acredited Laboratories = 16 and number of Volunteer Laboratories = 13

Test Methods
This is the first *Ustilago nuda* proficiency test to include Method 7-013b. Figure 4 shows that the median for methods 7-013a and 7-013b were very similar. For Lot 2, seven and six infected embryos respectively; and Lot 3, 13 and 13 infected embryos

respectively. Laboratories 11, 27 and 34 were excluded because they had very large outliers or used a modified method or did not provide test method used.

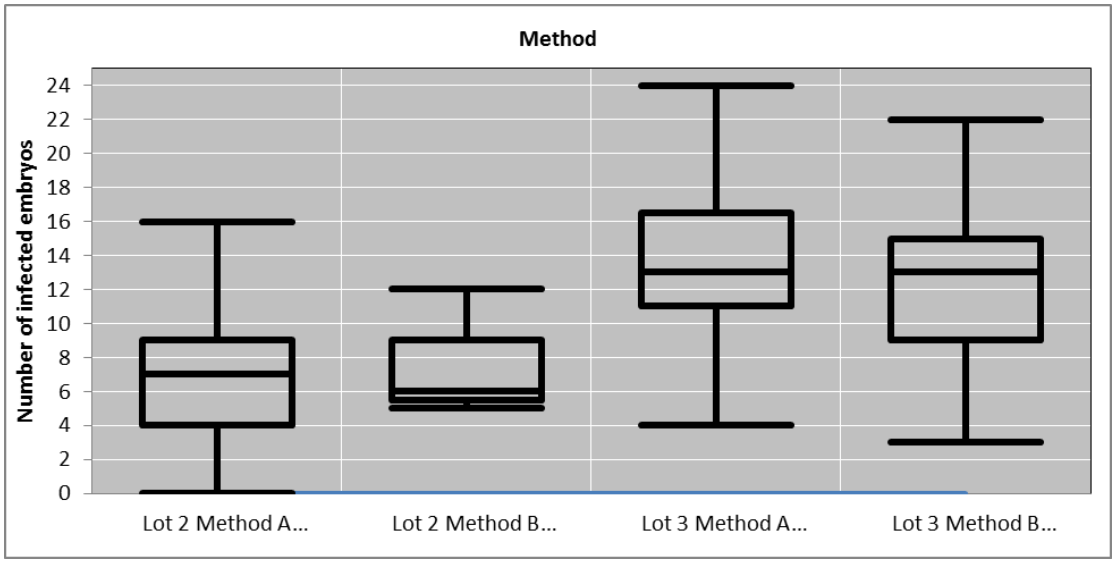


Figure 4 Comparison of laboratory results for Method 7-013a and 7-013b. (Excluding Lab 11, Lab 27 and Lab 34 results)

Conclusions

The majority of laboratories gave satisfactory results for all seed lots. There were still some issues with over estimation of infected embryos and false positives for some laboratories.

Analysis of data with low values should be given further consideration. As shown for Lot 2 a possible alternative is to consider data as binomial with possible outliers and estimate the proportion for the Binomial by the mean of the individual proportions and the p-value. It would be better to allow the possibility of greater variability (lower reproducibility), perhaps through the use of an over-dispersed version of the binomial distribution such as the beta-binomial. It should be possible to extend this method to robustly (i.e. not being affected by outliers) estimate the level of variability between labs. For cases such as lot one where the mean proportion is zero, more thought is required. This estimated proportion is treated as the true population proportion rather than a sample estimate of the proportion. Given that it is an estimate, then some allowance is needed for very low numbers of infected embryos apearing in some samples.

Introduction of a new equivalent method has shown no real differences between methods for contaminated lots.

The new rating system when applied to Lot 1 and Lot 3 provided reasonable differentiation between laboratories, but more work needs to be done to consider data with low values before using this rating model for infection levels like those found in Lot 2.