Technical Working Party on Testing Methods and Techniques   TWM/2/4

Second Session   **Original:** English
Virtual meeting, April 8 to 11, 2024   **Date:** March 12, 2024

REFERENCE COLLECTION MANAGEMENT USING MOLECULAR MARKERS: A NEW APPROACH BASED ON GENOMIC PREDICTION

Document prepared by experts from the United Kingdom

Disclaimer: this document does not represent UPOV policies or guidance

INTRODUCTION

1.   Currently UPOV Guidance describes two application models for the use of molecular markers in the examination of Distinctness, Uniformity and Stability (DUS), that have received a positive assessment (document TGP/15):

    a)   Characteristic-Specific Molecular Markers
    b)   Combining Phenotypic and Molecular Distances in the Management of Variety Collections

2.   Here we propose a new approach that is suitable when characteristic-specific markers are not available, and that has the potential to outperform model b).

3.   This development arises from INVITE, a collaborative project between European Examination Offices and researchers. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 817970.

Motivation

4.   Model a) is a useful approach for crops where there are highly heritable characteristics, such as some types of disease resistance. However, the utility will be limited to such characteristics, and not all crops include such characteristics in their Test Guidelines. For greater leverage in managing trials, it is beneficial to use more characteristics, such as under model b). A possible criticism of the current examples of model b) is that they are not targeted. Rather than compare molecular distance to individual characteristics, an overall phenotypic distance is used. This phenotypic distance is not used elsewhere in the process for assessing distinctness; characteristics are assessed one-by-one.

5.   We propose a new approach, which combines the benefit of a targeted linkage between markers and characteristics of model a), with the coverage over all characteristics of model b). In principle, we believe that this approach should outperform model b), and is more consistent with the UPOV approach.

The concept

6.   In this new approach, genomic prediction is used to predict the distinctness between a candidate and a reference variety in a particular characteristic. This information can then be used to manage the trial, by identifying reference varieties that are so likely to be distinct from the candidate that they do not need to be compared in the field.

7.   The information on which pairs of varieties are likely to be distinct for a given characteristic is then combined over all characteristics. In this way, the process benefits from distinctness only being required in one characteristic.

8.   We believe the new approach should outperform the current examples of model b) because:

- It should be possible to get better predictions for individual characteristics based on genomic predictions, compared to the typical correlations seen between phenotypic and molecular distances.
- It makes use of the principle that distinctness between a candidate and a reference variety is only required in one characteristic.

9.   The usefulness of this approach in practice (and for model b)) depends on the strength of the relationship between the phenotypic characteristics and the markers. This depends, at least in part, on the heritability of the characteristics. More characteristics with higher heritability will increase the number of reference varieties that can be identified as being distinct.

10.   It should be noted that, just like model b), where this approach is used to reduce the size of the growing trial, then the number of candidates plays a role in determining success. To eliminate reference varieties, they must be predicted as distinct from all the candidates. The greater the number of candidates, the fewer reference varieties that can be eliminated.

11.   Although we are not examining this in INVITE, genomic prediction also has the potential to improve the identification of similar varieties (model b example 2: see document TGP/15)

A note on Genomic Prediction

12.   Genomic prediction is a tool that uses genomic data to predict phenotypic outcomes. Although in INVITE, we concentrate on statistical models for this purpose, such as Genomic Best Linear Unbiased Prediction (GBLUP), there is no reason in principle why machine learning methods could not be used for this purpose. However, at least for one example crop that we have examined, more complex models such as the Bayesian alphabet cannot be successfully used due to the properties of the data set; we based our analysis on the most straightforward method, GBLUP.

Demonstration

13.   Within INVITE, this approach is being evaluated on four crops: maize, perennial ryegrass, soybean and wheat. We will show results for at least one crop.

Acknowledgments

14.   This work relied on collaboration with many partners, notably Karl Schmid (University of Hohenheim) and Dan Milbourne (Teagasc). We are grateful to the Examination Offices for supply of data and material. Thanks to CPVO and Naktuinbouw for useful discussions.

For more information: contact adrian.roberts@bioss.ac.uk

[End of document]