International Oat Nomenclature Committee Meeting

6 July 2021 / 13:00 UTC / Zoom Conference

Attendees

Martin Mascher (host), Rick Jellen, Victoria Blake, Matthias Herrmann, Yung-Fen Huang, Tim Langdon, Chengdao Li, Robert Park, Taner Sen, Charlene Wight, Manual Spannagl

Previous Topics

- Chromosome naming system was approved by the community. Charlene is maintaining the correspondence table.
- Minutes from 30Apr2021 were approved and are online at GrainGenes https://wheat.pw.usda.gov/GG3/oatnomenclature

The Oat Pan Genome.

- A nomenclature scheme is needed, since assemblies are nearly complete and annotation is beginning.
- Annotation will map transcript data to assemblies with pipelines to generate gene IDs.
- OT3098 (PepsiCo genome) wants to use the same nomenclature as PanOat.

Two examples were discussed.

1. AVESA.BEL.v1.4DG0792660.1
    a. AVESA - (<u>Ave</u>na <u>sa</u>tiva)
    b. BEL - germplasm abbreviation (i.e., Belinda)
    c. v1 - version 1
    d. 4D - chromosome
    e. G - gene
    f. 0792660 - gene number in position order
    g. .1 - isoform / splice variant

2. AVESA.00010a.r1.4DG0792660.1
    a. AVESA - (<u>Ave</u>na <u>sa</u>tiva)
    b. 00010a - germplasm code maintained at the lookup table at GrainGenes
        1. ('a' is the first annotation)
    c. r1 - release version 1
    d. 4D - chromosome
    e. G - gene
    f. 0792660 - gene number in position order
    g. .1 - isoform / splice variant

Discussion Notes:

Species abbreviations are in progress and are considering the following:

1) Should AVESA be used for all ‘natural’ hexaploids?
2) Species IDs have rules impugned by public repositories.

Discussion about human-readable germplasm abbreviation vs. a code to be maintained in a lookup table was perfunctory, as the decision was already made.

GrainGenes (Taner Sen) will maintain the lookup table that will contain the IDs, ploidy, sub-genera, etc..

Leading zeros in the germplasm code will be included for parsability.

The IONC will assign germplasm IDs.

Gene names; i.e., second level of ID, are assigned by homology or experimental evidence. There is a wheat initiative in progress for assigning the genes from the WGC to gene identifiers. (Spannagl)

Numberings across the chromosomes are independent from each other. If the identifiers are exactly the same, it doesn’t mean it’s the same gene. Establishing the ID’s of true orthologs is really a huge bioinformatics effort. Manual curation is necessary.

High/Low confidence will not be encoded in the identifiers, as the assignment of LC/HC is too fluid.

Follow-Ups

Jellen: species abbreviation list.

Jellen, Mascher, Sen: Google Doc for comment on the method to register genotypes.

Wight: Formalize correspondence table to post on the GG IONC page.

Invite Jizeng Jia to the IONC