

Guidelines for the annotation of information for marker assisted selection in wheat

FSOV Sam Blé project

Version 1.1 du 12 avril 2013

Authors: Claire Nédellec, Marion Ranoux, Pierre Sourdille, Dialekti Valsamou



This document presents the guidelines for the manual annotation of entities and relations involved in the representation of maker-assisted selection in wheat in the SAMblé project.

Référence

Marion Ranoux, Claire Nédellec, Emmanuelle Cariou-Pham, Robert Bossy, Claude Pope de Vallavieille, Dialekti Valsamou, Delphine Hourcade, Laurent Gueirero, Pierre Sourdille. [Validation de marqueurs liés à des gènes d'intérêt en vue de l'établissement d'une base de données pour la sélection assistée chez le blé tendre.](#) Synthèse des programmes de recherche FSOV (Fond de soutien à l'obtention végétale), Actes de la rencontre scientifique du 8 janvier 2015, 10 pages, Paris.

Outline

1. Annotation of entities	4
1.1 Entity types	4
1.2 Gene	4
1.2 Allele	4
1.3 Marker	4
1.4 Trait	4
1.4. Phenotype	5
1.5. Variety	5
1.6. AlleleSize	5
1.7. Type	5
1.8 Boundaries	6
1.9 Annotated section	6
2. Annotation of relationships	6
2.1 Binary relations	7
2.1.1 Relation marker_of_type.....	7
2.1.2 Relation marker_tags_gene.....	7
2.1.3 Relation gene_controls_trait	7
2.1.4 Relation variety_has_phenotype	7
2.1.5 Relation genes_allele_expresses_phenotype.....	7
2.1.6 Relation trait has value	7
2.1.7 Relation marker_alleleSize	8
2.2 Ternary relations	8
2.2.1 Relation marker_tags_gene_in_variety	8
2.2.2 Relation gene_expresses_phenotype_in_variety	8
2.2.3 Relation trait has phenotype in variety	8
2.2.4 Relation marker_size_in_variety	9
2. Annotation of entity groups	9

1. Annotation of entities

1.1 Entity types

1. Gene
2. Marker
3. Trait
4. Phenotype
5. Variety
6. AlleleSize
7. Type

1.2 Gene

The Gene entity type includes genes stricto sensu, but also proteins.

Examples

*leaf rust resistance gene **Lr51** and their use*

*PCR isolation of **LRIRGA1** alleles*

***ADP-glucose pyrophosphorylase** from wheat endosperm*

***High-molecular-weight glutenin subunits (HMW-GS)**, one class of seed storage proteins*

1.2 Allele

The type of entity Allele strictly refers to alleles.

Example

*only two alleles, photoperiod insensitive (**Ppd-D1a** and **Ppd-B1a**) and sensitive (**Ppd-D1b** and **Ppd-B1b**),*

1.3 Marker

Markers are designated by proper names or represent loci.

Examples

*the region between **RFLP** markers **ABC718** and **PSR567***

S19M93-140** completely cosegregates with **Yr5

1.4 Trait

The traits refer to the observable characters but do not include the value, the observation of the trait (the phenotype). It may include details of the species that expresses the trait (e.g. *stripe rust of wheat*), or of the pathogen (e.g. *race 5*). The trait does not include the conditions for the trait (e.g. *adult-plant*).

Examples

*In tests for **resistance to *P. triticina* race 5**, plants*

*wheat cultivars to provide **protection from WSM***

*High—temperature adult-plant **resistance to stripe rust of wheat**.*

1.4. Phenotype

The phenotype is the value of the trait.

Examples

cultivars that originally were resistant to leaf rust

Eight dwarf and semi-dwarf varieties, covering a range of genetic sources

The phenotype term may not include the term of the trait.

and highly-susceptible cultivar Wheaton

The phenotype must not be confused with the environmental factor.

Photoperiod has an important effect on plant growth

1.5. Variety

The variety entity is the proper name or a description of how the cultivar has been obtained. La variété peut être un nom propre où une description de l'obtention du cultivar. The name of the species is not part of the variety name.

Examples

analysis of Thatcher near-isogenic lines with wheat leaf rust resistance genes

Mv Summa and Mv Emma winter wheat varieties

the spring wheat variety Bobwhite SH 98 26 Triticum aestivum L.

*Near-isogenic lines for Lr48 (CSP44/5*LB) and Lr49 (VL404/5*LB) were isolated*

1.6. AlleleSize

The AlleleSize entity represents the size in base pairs (*bp*) of the allele for a given marker. The entity includes the words *Fragment* and *DNA fragment* when presents.

Examples

absence of the 275-bp marker in the F2 population

590- bp DNA fragment was present in all the 22 wheat cultivars

the Xgwm633 marker revealed a polymorphic fragment of 200 bp on the resistant wheat line 99419G4-1A/1-1

110-bp amplified product

When the restriction enzyme is mentioned, it is also annotated with the allele length.

Examples

absence of the 11-kb BamHI-fragment or the 8-kb HindIII-fragment

A band (11-kbp EcoRI, 8-kbp BamHI and 2-kbp HindIII fragments) was absent

1.7. Type

The Type entity represents the type of the marker.

Examples

the two flanking SSR markers identified in this study

the PCR primers J15 for the WSMV resistant gene

The indication of the proximity of the markers with the genes is not part of the type. For example, the word *flanking* in *recombination frequencies between the two flanking markers Xgwm389 and Xgwm566*.

1.8 Boundaries

Exclusion

The information that are not part of the entity type must not be included in the entity span.

Do not include the type of the entity itself.

SSR marker

Do not include irrelevant precisions.

*resistant gene in CO960293-2 as Wsm2
targeted 241-bp fragment
wheat cultivar Ron L*

Discontinuous annotation

Distinct entities must be annotated separately.

resistant to both WSMV and Triticum mosaic virus

1.9 Annotated section

The section of the text that have been fully annotated should be annotated by the Relevant entity, so that they can be distinguished from the non-annotated or partly annotated sections.

2. Annotation of relationships

Relationships are oriented. They connect entities. They have a definite meaning. Each entity has its defined place as an argument of the relationship. For example, the *marker_of_type* relationship links a marker and its type. The first argument must be an entity of maker type and the second argument is an entity of type type.

Binary relationships link pairs of entities, ternary relationships link triplets of entities. Ternary relationships should be preferred to binary relationships whenever possible.

2.1 Binary relations

Binary relations are ,


1. marker_of_type
2. marker_tags_gene
3. gene_controls_trait
4. variety_has_phenotype
5. genes_allele_expresses_phenotype
6. trait has value
7. marker_alleleSize

2.1.1 Relation marker_of_type

The relation links the marker to its type.

marker_of_type (marker, type)

the region between RFLP markers ABC718 and PSR567




2.1.2 Relation marker_tags_gene

The relation links the marker to the gene.

marker_tags_gene (marker, gene)

SI9M93-140 completely cosegregates with Yr5




2.1.3 Relation gene_controls_trait

The relation links the gene to the trait that it controls.

gene_controls_trait (gene, trait)

Lr21 leaf rust resistance gene

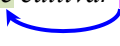


2.1.4 Relation variety_has_phenotype

The relation links the variety to the phenotype that it shows.

variety_has_phenotype (variety, phenotype)

highly-susceptible cultivar Wheaton




2.1.5 Relation genes_allele_expresses_phenotype

The relation links the allele (the gene if the allele is not named) to the phenotype that it determines.

genes_allele_expresses_phenotype (gene, phenotype)

The gene Wsm1, transferred to wheat from intermediate wheatgrass confers resistance to WSMV



2.1.6 Relation trait has value

The relation links the trait to the observed phenotype.

trait_has_value (trait, phenotype)

six wheat genotypes were compared for their responses to WSMV infection under growth chamber conditions. The three resistant genotypes,

2.1.7 Relation marker_alleleSize

The relation links the marker to its size.

marker_alleleSize (marker, alleleSize)

GWM296 amplified a polymorphic pattern displaying a DNA fragment (145 bp) specific to the resistant germplasm

2.2 Ternary relations

There are four ternary relations.

1. marker_tags_gene_in_variety
2. gene_expresses_phenotype_in_variety
3. trait has phenotype in variety
4. Relation marker_size_in_variety

2.2.1 Relation marker_tags_gene_in_variety

The relation marker_tags_gene_in_variety links the marker to the gene for a given variety.

marker_tags_gene_in_variety (marker, gene, variety)

RL6005, which carries Lr16, differed from these lines only for WMC764.

2.2.2 Relation gene_expresses_phenotype_in_variety

The relation gene_expresses_phenotype_in_variety links the allele of the gene to the phenotype for a given variety.

gene_expresses_phenotype_in_variety (gene, phenotype, variety)

The resistance to WSMV in CO960293-2 was found to be governed by the single dominant gene Wsm2

2.2.3 Relation trait has phenotype in variety

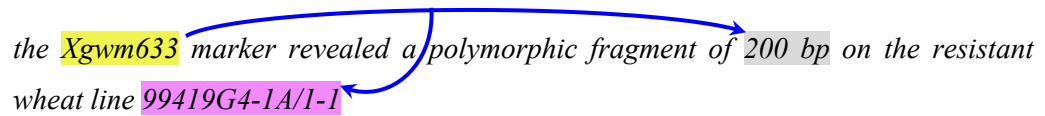
The relation trait_has_phenotype_in_variety links the trait to the observed phenotype for a given variety.

trait_has_phenotype_in_variety (trait, phenotype, variety)

2.2.4 Relation marker_size_in_variety

The relation `marker_size_in_variety` links the marker to its size for a given variety.
`marker_size_in_variety` (marker, alleleSize, variety)

the **Xgwm633** marker revealed a polymorphic fragment of 200 bp on the resistant wheat line **99419G4-1A/1-1**

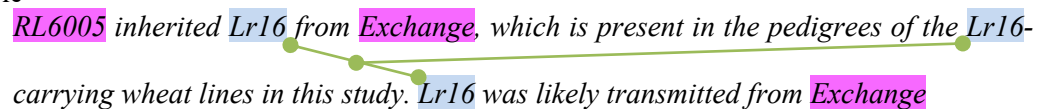


2. Annotation of entity groups

The entity groups are used to annotate entities that are equivalent. This avoids having to repeat the annotation of the same relationship between several occurrences of the same entity. It is possible to create groups for each of the entity types. The most common group is that of genes.

Example

RL6005 inherited **Lr16** from **Exchange**, which is present in the pedigrees of the **Lr16**-carrying wheat lines in this study. **Lr16** was likely transmitted from **Exchange**

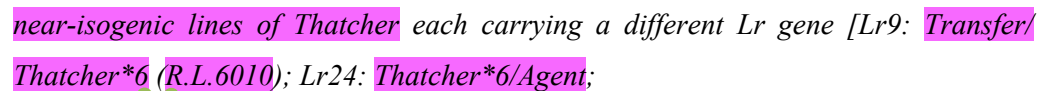


The notion of group also allows to annotate synonyms.

It is not useful to annotate groups if there are no relationships involving the entities of the group, except to annotate synonyms of different forms.

Example

near-isogenic lines of Thatcher each carrying a different Lr gene [Lr9: **Transfer/Thatcher*6** (R.L.6010); Lr24: **Thatcher*6/Agent**;



This gene, temporarily named **LrF7**, has been designated **Lr51**

