

Stephan Weise

Phenotypic data in EURISCO

Status and plans

Third Meeting of the EURISCO Advisory Committee
7th July 2021



Phenotypic data search in EURISCO

Filter C&E data by genus

Genera * Brassica Capsicum Chondrilla Cicerbita Cucumis Eruca Ixeridium Linum Lupinus Mycelis

Allium Hordeum Lactuca

Apply Reset

Lactuca	105,02
Solanum	77,663
Capsicum	50,736
Triticum	37,301
Hordeum	32,852
Brassica	27,355
Spinacia	17,913
Cucumis	17,460
Pisum	17,233
Linum	14,354

Filter C&E data by species and traits

Genus * Lactuca

Species * Lactuca aculeata Boiss. Lactuca altaica Fish. & Mey. Lactuca biennis (Moench) Fern. Lactuca homblei De Wild. Lactuca raddeana Maxim Lactuca saligna L. Lactuca sativa L. Lactuca sativa x serriola Lactuca serriola L. Lactuca tatarica (L.) C. A. Mey. Lactuca canadensis L. Lactuca dregeana DC. Lactuca georgica L. Lactuca perennis L. Lactuca indica L. Lactuca quercina L.

Traits * Leaf color intensity ((3=light, 5=medium, 7=dark[...])) Leaf margin undulation (At harvest maturity[...]) Leaf shape ((1=narrow elliptic, 2=el[...]) Leaf shape ((1=round, 2=ovate, 3=obov[...]) Leaf vein prickles ((1=not present, 9=present[...]) Leaf vein prickles (-[...]) Leaf venation (At harvest maturity (1= n[...]) Nasonovia ribisnigri (Resistance to Nasonovia r[...]) Nitrate content (Mean nitrate content of t[...]) Pemphigus hirsarius ((1=very resistant, 2=resil[...])

Apply Reset

Filter C&E data by experiment

The report below lists all experiments, which contain characterisation & evaluation (C&E) data. Please use the search bar below to define filters.

Experiment Start Year between 1967 and 2012

1 - 10 of 782

Experiment Description	Dataset Remark	Experiment Start Year	Experiment End Year	Details
Sowing date = February 2, Planting date = April 17, IVT glasshouse XII, heated, soil culture, 2 stems, 4 plants per field, collection no. 567-659, experimentalist H. Roelofs and G. Pet, standard = Bruinsma Wonder	Test data CGN	1980	-	contained traits
Sowing date February 16, Planting date April 8, IVT glasshouse XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 444-543, experimentalist L. de Groot and G. Pet, standard = Bruinsma Wonder				
Sowing date = March 15, Planting date = April 26, IVT glasshouse XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 660-762, experimentalist L. de Groot and G. Pet, standard = Bruinsma Wonder				
Sowing date = February 28, Planting date = April 13, IVT glasshouse XII-IX, heated, soil culture, 2 stems, 5 plants per field, collection no. 763-869, experimentalists L. de Groot and G. Pet, standard = Bruinsma Wonder				
Sowing date = February 24, Planting date = April 18, IVT glasshouse no. XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 871-934, experimentalists L. de Groot and G. Pet, standard = Bruinsma Wonder				
Sowing date = March 11, Planting date = April 26, IVT glasshouse XII, heated, soil culture, 2 stems, 5 plants per field, collection no. 935-981, experimentalist L. de Groot and G. Pet, standard = Bruinsma Wonder				
Sowing date = March 13, Planting date = May 1, IVT glasshouse II-I, heated, soil culture, 2 stems, 5 plants per field, collection no. 982-1021, experimentalist G. Pet, standard = Bruinsma Wonder				
Sowing date = March 20, Planting date = April 28, IVT glasshouse no. II-II, soil culture, 1 stem, 5 plants per field, collection no. 1476-1574, experimentalist G. Pet, standard = Sonatine				
Sowing date = January 31, Planting date = March 31, IVT Glasshouse no. 12-7, heated, soil culture, 2 stems, 5 plants per field, collection no. 33-66, experimentalist G. Pet, Standard = Claessee				
Sowing date = January 29, Planting date = March 28, IVT glasshouse no. 12-5, heated, soil culture, 2 stems, 5 plants per field, collection no. 1-111, experimentalist G. Pet, standard = Claessee	Test data CGN	1979	-	contained traits

1 - 10 of 782

0.03 s

Wizard-based searches for

- Genus
- Species and trait
- Experiment
- Trait

Traits in selected experiment

1 - 10 of 26

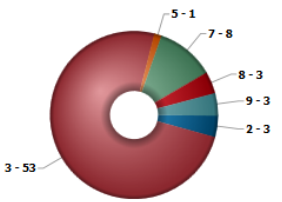
Trait Name	Trait Remark	Trait Method	Details
Fruit corrugation	-	(0=smooth, 3=slightly corrugated, 5=medium, 7=corrugated, 9=very corrugated)	scores
Fruit attitude	-	Bruinsma Wonder=7 (1=very drooping, 3=drooping, 5=horizontal, 7=semi-erect, 9=erect)	scores
Flower attitude	-	Bruinsma Wonder=7 (1=very drooping, 3=drooping, 5=horizontal, 7=semi-erect, 9=erect)	scores
Mature fruit color	-	(A=dark red, B=light r, C=orange, D=salmon, E=canary, F=sulphur, G=green, I=brown, J=light orange, K=white, a=b=both in one fruit)	scores
Tobacco mosaic virus	-	determined at natural infection (0=no symptoms, +=symptoms present)	scores
Stem anthocyanin content	-	Bruinsma Wonder=3 (0=absent, 1=very little, 3=little, 5=medium, 7=much, 9=very much)	scores
Fruit ribbing	-	(0=absent, 1=very little, ..., 9=very high)	scores
Flower color	-	(A=white, B=filly-white, C=light green, D=light purple, E=dark purple, F=yellow, G=white/anthocyanin)	scores
Fruit outwall thickness	-	Measurement, 9=9mm or more.	scores
Fruit cracking tendency	-	(1=none, 3=slight, 5=medium, 7=medium to severe, 9=severe)	scores

1 - 10 of 26

0.12 s

Phenotypic data search in EURISCO

Trait details



Descriptive statistics

Trait Name	Minimum	Maximum	Average	Stddev	Variance	First Quartile	Median	Third Quartile
Fruit attitude	2	9	3.9	1.93	3.72	3	3	3

Experiment description: Sowing date = February 2, Planting date = April 17, I/VT glasshouse XII, heated, soil culture, 2 stems, 4 plants per field, collection no. 567-659, experimenter H. Roelofsen and G. Pet, standard = Bruinsma Wonder

Trait name: Fruit attitude

Additional filters

Genus:

Origin Country:

Accession scores for selected trait

Q Go Rows 10 Actions

1 - 10 of 71 >

NICODE	INSTCODE	GENUS	ACCENUMB	Score	Score Link	Origin Country	Biological Status	Details
NLD	NLD037	Capsicum	CGN16913	3	-	Germany	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16914	3	-	Netherlands	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16916	8	-	Israel	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16917	7	-	Israel	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16918	3	-	Hungary	Traditional cultivar/landrace	Accession details
NLD	NLD037	Capsicum	CGN16919	3	-	Hungary	Advanced or improved cultivar (conventional breeding methods)	Accession details
NLD	NLD037	Capsicum	CGN16920	7	-	-	-	Accession details
NLD	NLD037	Capsicum	CGN16904	9	-	-	-	Accession details
NLD	NLD037	Capsicum	CGN16905	3	-	-	Traditional cultivar/landrace	Accession details
NLD	NLD037	Capsicum	CGN16906	3	-	-	Advanced or improved cultivar (conventional breeding methods)	Accession details

1 - 10 of 71 >
0.02 s

Wizard-based searches for

- Genus
- Species and trait
- Experiment
- Trait

Refine result

- Sort
- Filter
- Download
- Chart

Dealing with phenotypic data: Great diversity

- Phenotypic data
 - Determines value of germplasm for breeding and research
 - Crop-specific traits and methods
 - Many historical datasets
 - Usually no data from high throughput phenotyping
 - Data has to be aggregated or exchanged between organisations

Lots of “standards” to express traits

- Different trait names/synonyms
- Different rating scales (nominal, ordinal, metric)

Different amounts of meta information

- When, where, how, by whom?
- Experiment set-up, treatment etc.

Different means of data management

- DBMS, flat files, mainly Excel files

Dealing with phenotypic data : Existing situation

Methods and Descriptors

- Crop-specific definitions of traits, methods etc. like IPGRI descriptor lists
- Often used in parts only and adapted to organisational needs

Exchange Formats

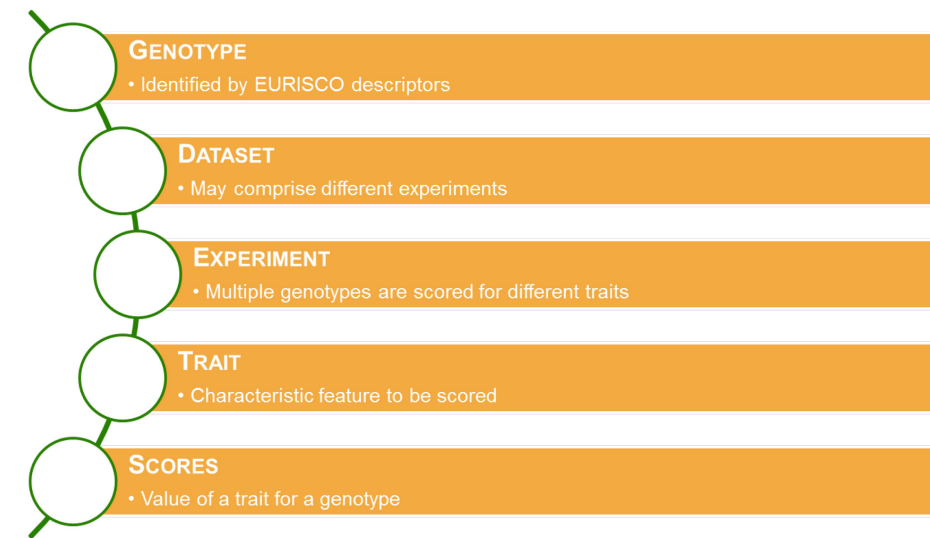
- E.g. Darwin Core germplasm extension (DwC-germplasm; Endresen et al. 2009)
- Great for computer scientists
- Difficult to handle for genebank curators

Ontologies

- Help to structure the (phenotypic) world
- Improve interoperability of data
- e.g. Crop Ontology (Arnaud et al. 2012)

Dealing with phenotypic data: Current approach

- Data standardisation
 - About 600 germplasm collections in Europe, around 400 in EURISCO
 - No standardisation of trait, scale or experimental design
 - Pragmatic approach: Import of existing data as-is to reach critical mass
- Data exchange
 - Only standardisation of exchange format
 - As simple as possible
 - As few fields as possible
 - “minimum consensus”
- Data management
 - Highly abstracted, following the single-observation concept (van Hintum et al. 1992)
 - Omitting fine-grained metadata



Dealing with phenotypic data: Data overview

- Extension available since 2016
- 2,680,511 records
- 90,569 accs. with phenotypic data
- 17 countries
- 68 phenotypic datasets
- 3,862 experiments
- 9,415 traits
- Increasingly accepted as repository, but limited comparability

Filter C&E data by trait

The report below lists the definitions of all phenotypic traits, which are currently available in EURISCO. Please use the search bar below to define filters.

Rows 10

1 - 10 of 61

Trait Name	Trait Method
Flowering time	Count days after 1 April when >50% plants show inflorescence emergence, 999=not flowering during experiment
Flowering time end	(3=early, 7=late)
Flowering time	Number of days between the date of sowing and the date of appearance of the first flower head
Flowering time begin	Days after sowing when 50% of plants have opened the first flower(s)
Flowering time	Count days after 1 September when >50% plants show inflorescence emergence, 999=not flowering during experiment
Flowering time	No treatment. Count days from planting to corolla 1st flower visible (1=<41. 2=41-60. 3=61-80. ... 8=161-180. 9=>180)
Flowering time	Count days to 10% of flowers have opened after sowing
Flowering time	count days after 1 May when 50% of florets have opened on 3 flowers
Flowering time begin	(3=early, 7=late)
Flowering time begin	Count the days from 25/5 to 50% of plants in flower

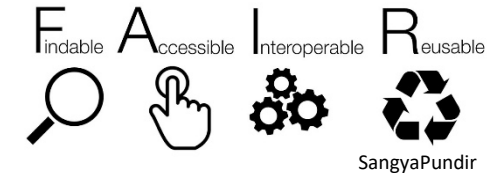
1 - 10 of 61

0.06 s

as of 2021-06-23

Dealing with phenotypic data: Towards FAIR data

- Data harmonisation
 - Experiment set-up, treatment etc.
 - Reach MIAPPE-compliance (Krajewski et al. 2015)
- Better structuring
 - Traits/methods/scales
 - Development of common vocabularies/approaches
 - Improve comparability
 - Mapping onto ontology terms
 - Ontology of choice: Crop Ontology (Arnaud et al. 2012)
 - Crux: Sustainability of ontologies
- Provide training + helpdesk
- Additional activities together with various partners, e.g. AGENT



AGENT as a blueprint

- Current limitations
 - EURISCO data exchange format represents a „minimum consensus“
 - Difficult to compile files manually
 - Very limited reproducibility and comparability
- AGENT approach
 - Simplification of data collection → one column per trait to support manual recording
 - Distinction in two types of data

Two “types” of phenotypic data

- Historical phenotypic data
 - Existing data, e.g. from seed multiplication
 - Collected over a long period of time
 - Relatively few metadata available
 - Focus on simple traits
- New experimental data
 - Desirable to describe the experiments as comprehensive as possible
 - Harmonisation of data of the various project partners
 - Traits
 - Methods
 - Experimental set-ups

Two “types” of phenotypic data (cont.)

- Historical phenotypic data
 - Simplified EURISCO-EVA templates should be sufficient
 - Already in use
- New experimental data
 - Templates presently used in EURISCO do not allow to collect additional metadata
 - Only very little information can be represented
 - Focussing on providers managing data in database management systems
 - Strive for increased user-friendliness
 - Consider more extensive metadata
 - Allow full MIAPPE compliance
 - Under development

Separate template for new data?

- Data to be collected
 - Actual observation data → relatively easy to handle
 - Metadata → can be compiled during a datathon
 - Experimental set-up (start and end, plot length and width, number of rows, distance between rows, seeds/m², experimental design, ...)
 - Treatments (herbicide, fungicide, insecticide, fertiliser, growth regulator, ...)
 - Experimental location (location/site name, coordinates, average values of precipitation, temperature, soil type, ...)
 - Soil information (N, pH, K₂O, P₂O₅, Mg,)
 - Traits
 - Definition (name, description, unit, type, method, ontology term)
 - Range of values (rating values, allowed min/max values)



M. Grau / IPK

Thank you for your attention