

INCREASE

Intelligent Collections of Food Legumes Genetic Resources for
European Agrofood Systems

Phenotyping by citizen scientists in INCREASE

PRO-GRACE-EMPHASIS policy symposium and workshop

Valerio Di Vittori, Adriano Mancini, Roberto Papa
r.papa@univpm.it Università Politecnica delle Marche
UNIVPM – D3A & VRAI



UNIVERSITÀ
POLITECNICA
DELLE MARCHE



CSE summary

<https://www.pulsesincrease.eu/experiment/app>



Download of the INCREASE CSE App (INCREASE CSA) and citizens registration



More than 1,000 common bean landraces in INCREASE CSE



Each participant receives five different landraces + one control variety, common among all citizens

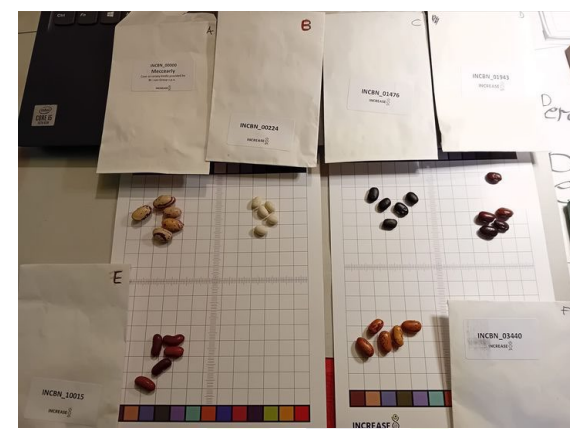


Citizens grow and evaluate the common bean CSE lines in plots or pots in fields, gardens, terraces, balconies, etc...

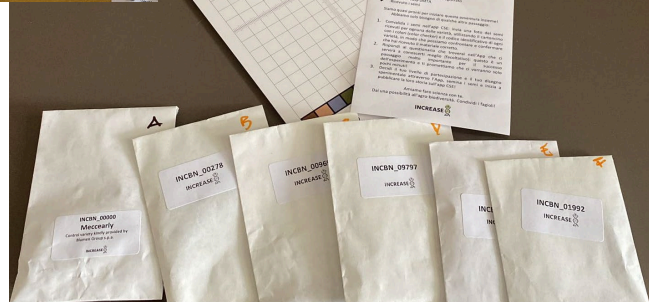
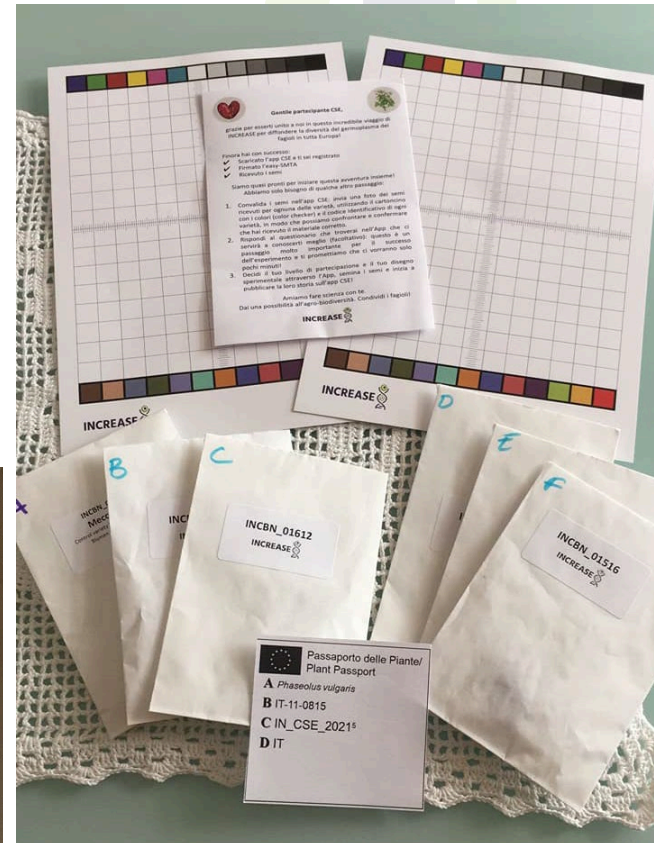
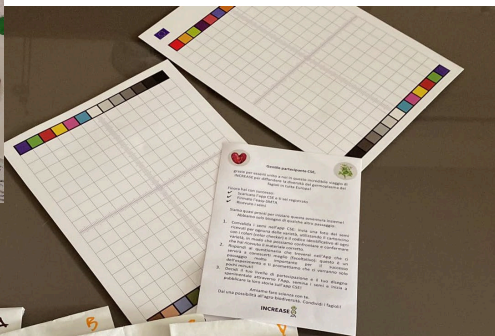
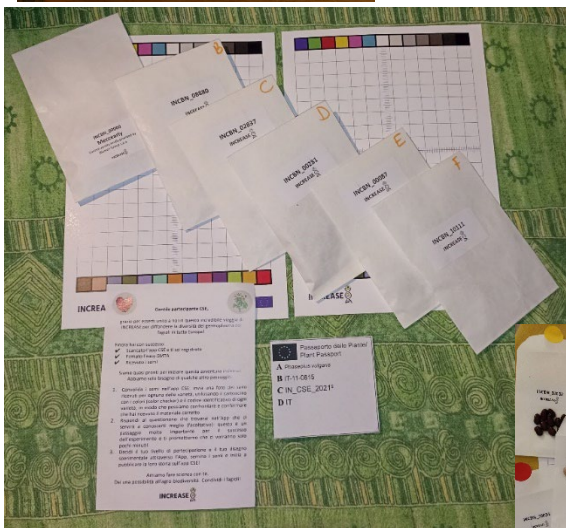
All interested citizens can voluntarily contribute and test a decentralised approach to seed conservation, multiplication and sharing in order to conserve and valorize agro-biodiversity



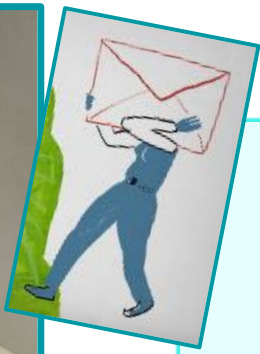
Through the App, citizens collect and record information on the common bean CSE lines they are nurturing, and share their experiences, uses, photos, opinions



Each citizen receives a postal parcel with common bean lines and the material to start the experiment and collect data



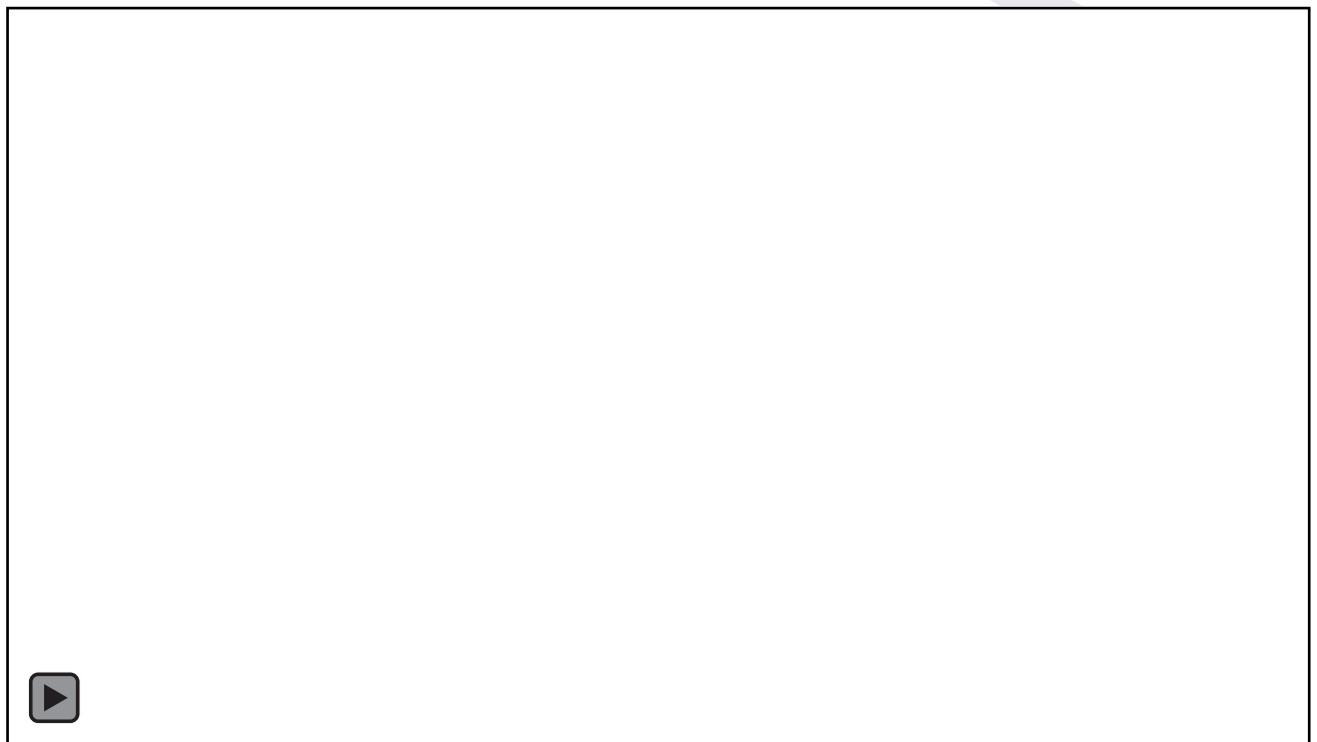
Images from Facebook groups, by EU citizens of Round 3 and Round 4



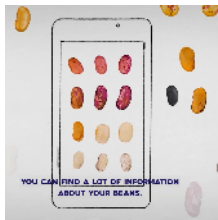
Every year, from the end of February/beginning of March, every registered citizen receives:
6 bags of common bean seeds
(5 landraces + the control variety)
from 5/6 seeds to 10 seeds



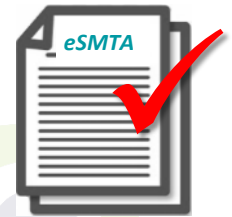
- ✓ **2 Color Checkers**
- ✓ **Welcome letter**
- ✓ **Seed Passport**
- ✓ **6 different seed bags** (INCREASE codes)



The **INCREASE CSA**; a dedicated App



- **Registration and acceptance of the Easy-SMTA**



- **Take Photos and Videos** to document common bean lines cultivation, in every moment of the growing season (and post harvesting)

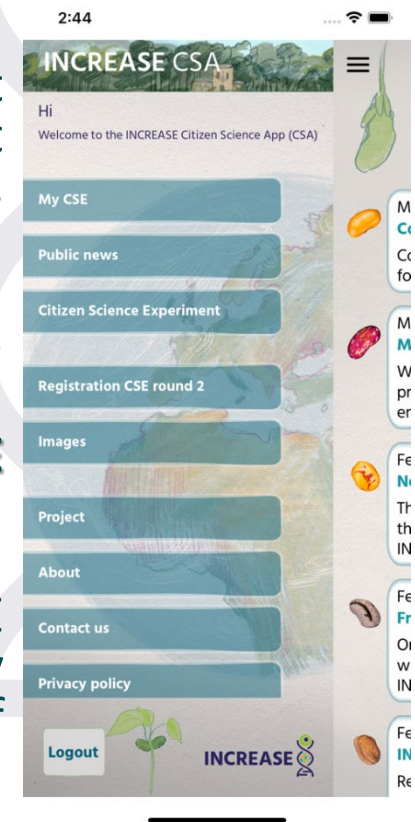
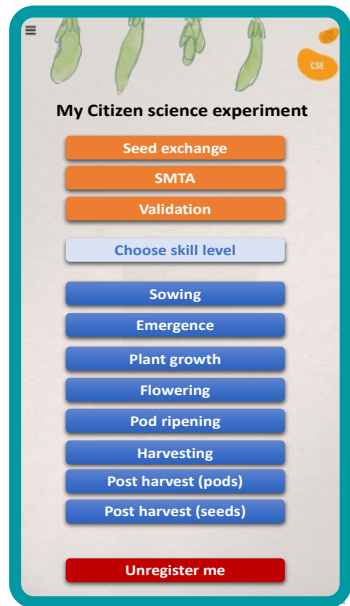


- Collect **Data and information** to document common bean lines cultivation, in every moment of the growing season (and post harvesting), following simple *protocols and tutorials*

- Request and receive information and assistance

- **EXCHANGE seeds with all EU INCREASE CSE participants**

- Keep updated on the CSE, on the INCREASE project activities and initiatives and play a key role in the decentralized conservation of legumes PGR

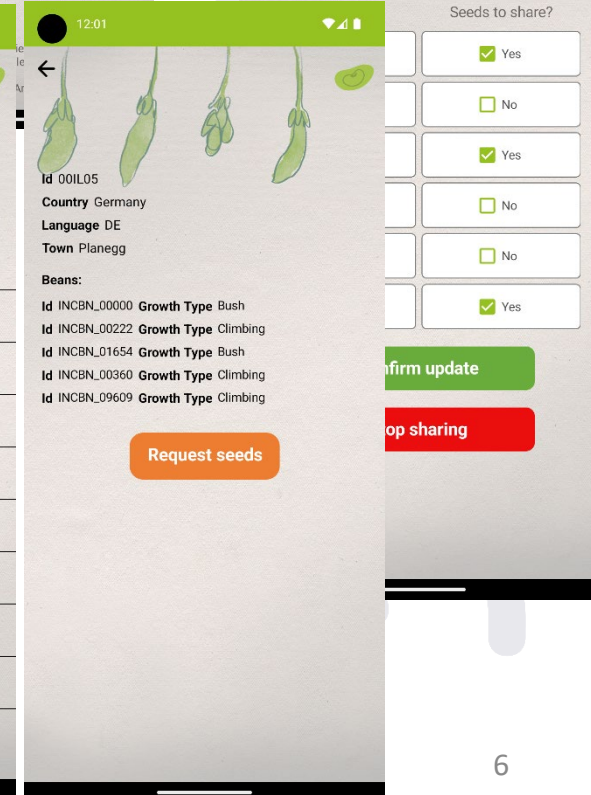
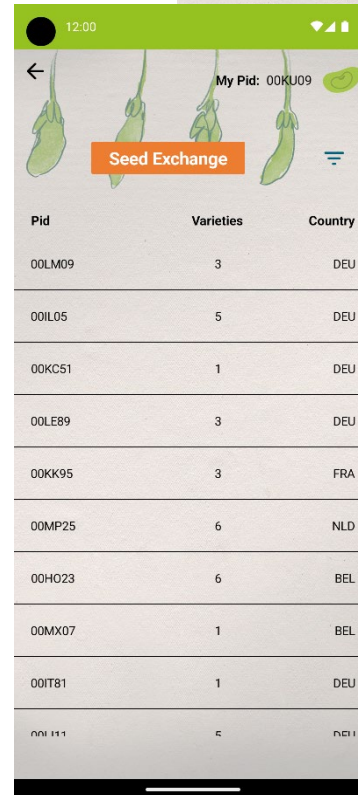
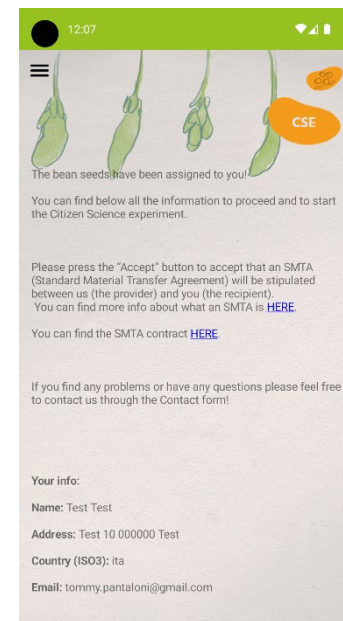


Easy SMTA

- Before participating in the CSE, users are asked to confirm their data. At this time a new FAO user is created if necessary, generating a PID.
- Before being sent their seeds, users have to accept the SMTA directly in the app.
- When the SMTA is signed, new DOIs are also generated.

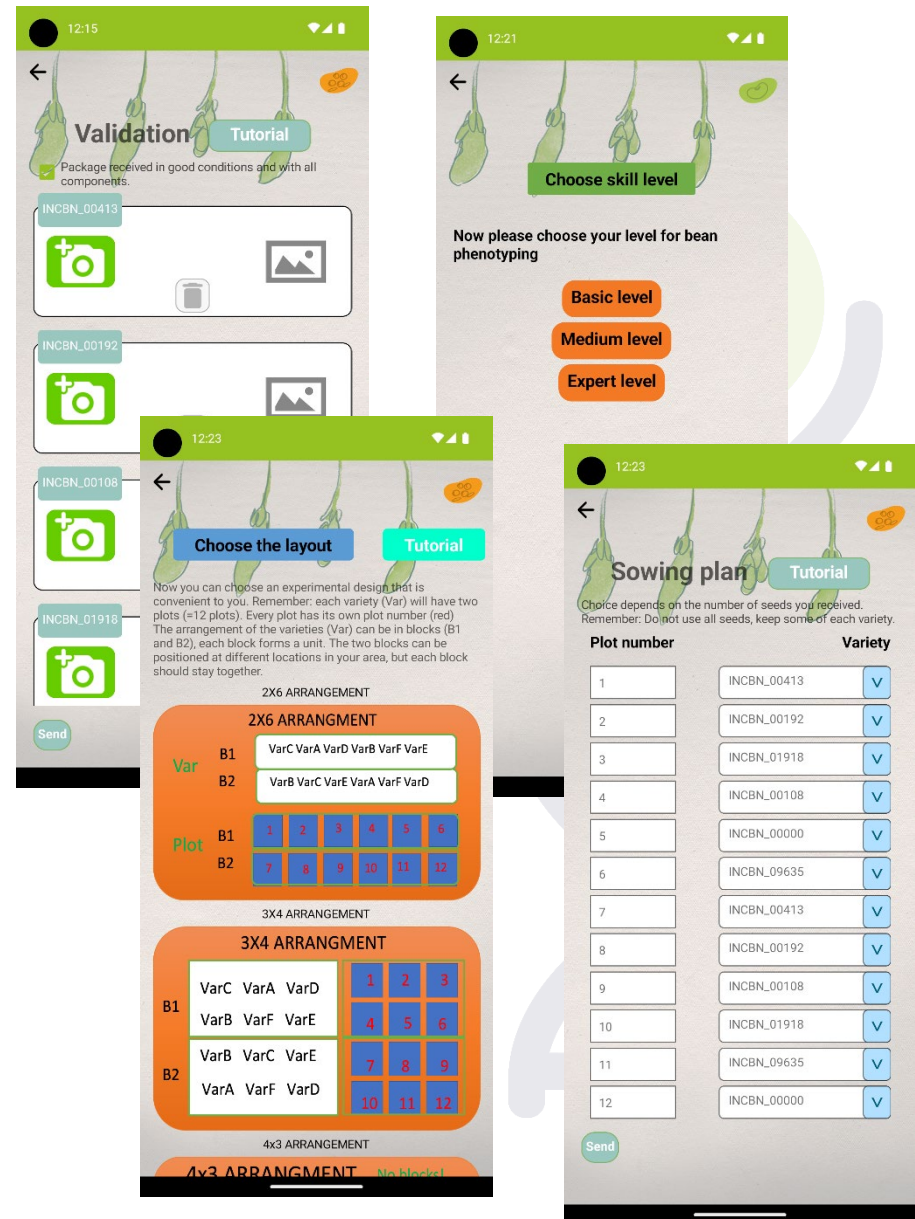
Seed Exchange

- Users can make available for seed exchange their leftover seeds
- Users can see a list of everyone that has seeds available for exchange
- From this list, user can request seeds directly to other users, without any intervention from our part
- Users can accept or refuse requests they receive
- SMTAs are signed and new DOIs are generated
- Returning users can also just sow their leftover seeds without engaging in seed exchange



Validation and Sowing

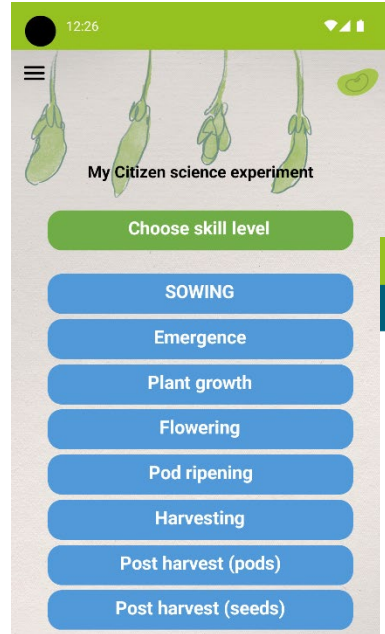
- Users have to validate the beans they receive. This is done by taking photos of everything they received.
- A skill level has to be selected: **Easy, Medium or Expert**. This can be changed at any time.
- Also a sowing layout must be selected.
- Finally a sowing plan is set, assigning a specific bean variety to each plot.



Phenotyping

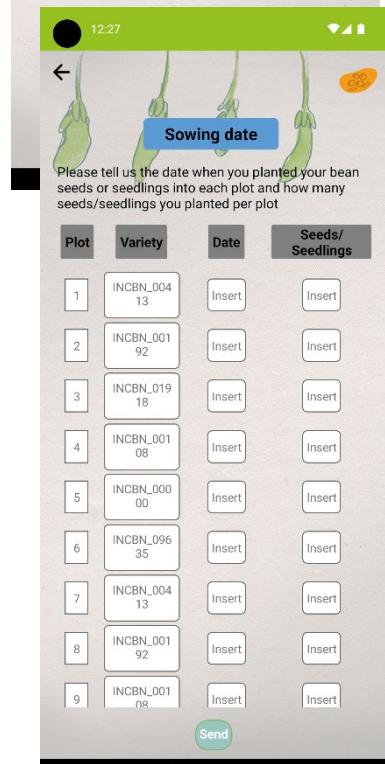
Easy	Medium (additional traits)	Expert (additional traits)
------	----------------------------	----------------------------

-Sowing date	-Plant emergence (date and number of plants)	-Hypocotyl pigmentation
-Flowering time	-Leaf shape	-Leaf colour; anthocyanin and chlorophyll
-Days to harvest	-Flower colour	-Stem diameter
-Total seeds weight	-Pod setting	-Maximum flowering
-Number of seeds	-Pod shape (curvature)	-Maximum pod setting
-Plant death	-Pod colour	-Fresh pod cross section
	-Growth habit and plant determinacy	-Plant health
	-Fibers presence in the pod	-Number of pods
	-Pod length and width	-Seed brilliance
	-Seed colour and pattern	-Seed length
	-Seed shape	-Seed width
		-Seed height



Bean flowers are about 1 cm long and have distinct upper and lower lobes. The flower position is at the node. Please score the date, when you see in your plot the first fully open flower. You have two possibilities:

1. Please take an image of the flower by placing the colour checker behind the flower. Please make sure that the scale and the colour reference is visible. When uploading the image, the upload date is automatically recorded by the App as flowering date. This is the preferred way.
2. If the image is older than the upload date or if you could not take an image but recorded the date of flowering manually, you have the possibility to tell us the flowering date directly. Please use that option only if the first possibility is not working for you.



CSE Instruction and tutorials

Flower colour

Please record the colour of the flowers (standard and wings) in the app.

Bean flowers have two different types of petals (see image below). The standard is the upper part and the two wings are forming the lower part. Please record the colour of both types when your plants have a decent number of flowers and chose a representative flower of a well-developed plant from each plot.

Note that often standard and wings have the same colour. However, there are bean varieties where standard and wing colour differ.

1. **Please take an image of the flower by placing the colour checker behind the flower.** Please make sure, the scale and the colour reference is visible (see picture and video).
2. **Please tell us your score for the flower colour of standard and wings using the indicated colour classification number.**

1 - Whitish



2 - Pink



3 - Purple



The common bean **INCREASE CSE lines**



1.126

Common bean landraces
(*Phaseolus vulgaris*)



+

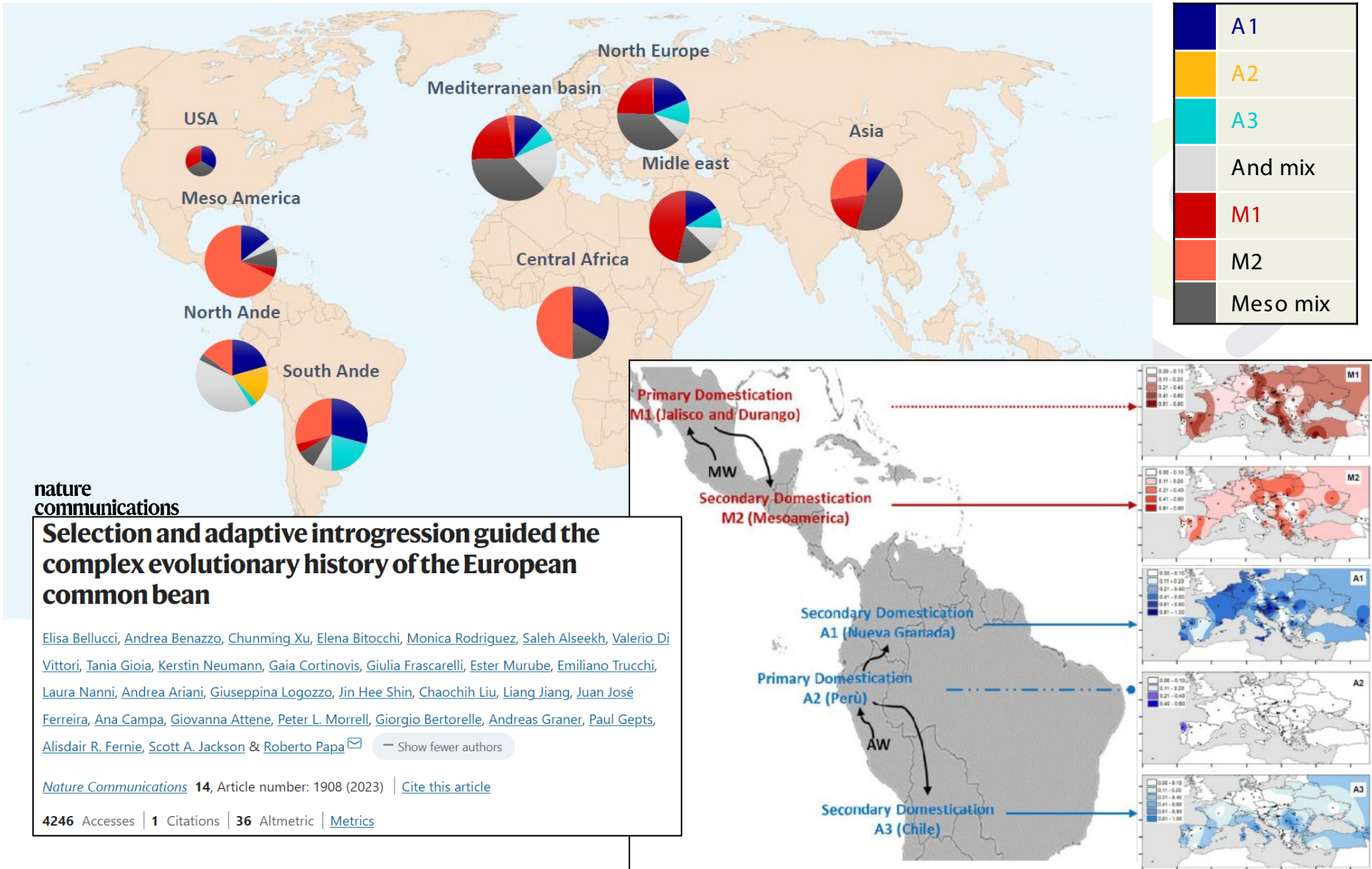
1 Control variety

Meccearly

early flowering variety, bush type
kindly provided by **Blumen Group** s.p.a.

- Different growth habit (bush, about 25% or indeterminate, about 75%)
- With passport data and representative of more than 40 countries worldwide
- Representative of the main gene pools and races of common bean
- **DOIs** and INCREASE codes assigned
- Included and analysed in the different Intelligent collections or in subsets for specific tasks and activities (eg. *Tcore* about 28%)





nature communications

Selection and adaptive introgression guided the complex evolutionary history of the European common bean

Elisa Bellucci, Andrea Benazzo, Chunming Xu, Elena Bitocchi, Monica Rodriguez, Saleh Alseekh, Valerio Di Vittori, Tania Gioia, Kerstin Neumann, Gaia Cortinovis, Giulia Frascarelli, Ester Murube, Emiliano Trucchi, Laura Nanni, Andrea Ariani, Giuseppina Logozzo, Jin Hee Shin, Chaochih Liu, Liang Jiang, Juan José Ferreira, Ana Campa, Giovanna Attene, Peter L. Morrell, Giorgio Bertorelle, Andreas Graner, Paul Gepts, Alisdair R. Fernie, Scott A. Jackson & Roberto Papa [Show fewer authors](#)

Nature Communications **14**, Article number: 1908 (2023) | [Cite this article](#)

4246 Accesses | 1 Citations | 36 Altmetric | [Metrics](#)

Genotypic analysis

Plant material gene-pool composition

In progress...
Completion of sequence
data collection

GBS data for 968

 CSE lines

50,155 SNPs
Biallelic
Polymorphic
Max Missingness 50%
MAF 1%

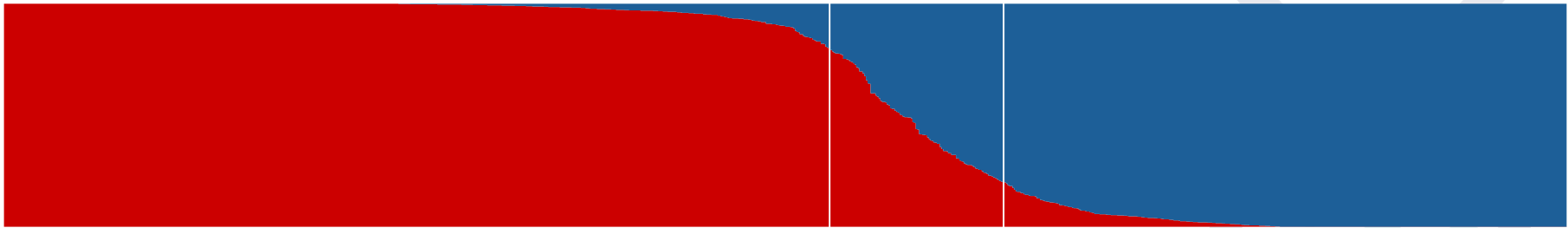
250k thinning

1,770 SNPs

→

Admixture
K2

0.8
membership
threshold






512
Mesoamerican

37,963 SNPs
Biallelic
Polymorphic
Max Missingness 50%
MAF 1%

250k thinning





1,682 SNPs

MESO Admixture
genepools
0.8 membership threshold

	M1 - 162
	M2 - 146
	Meso mix - 204

107
Mix

AND Admixture
genepools
0.8 membership
threshold

	A1 - 142
	A2 - 4
	A3 - 54
	And mix - 149

349
Andean

34,580 SNPs
Biallelic
Polymorphic
Max Missingness 50%
MAF 1%

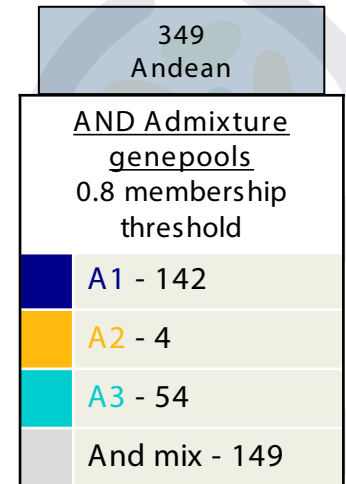
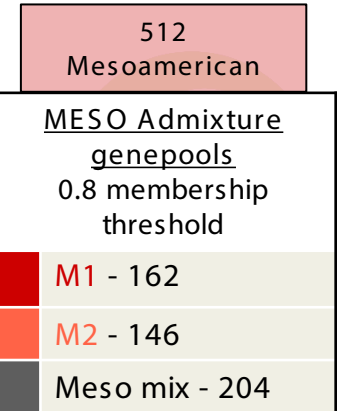
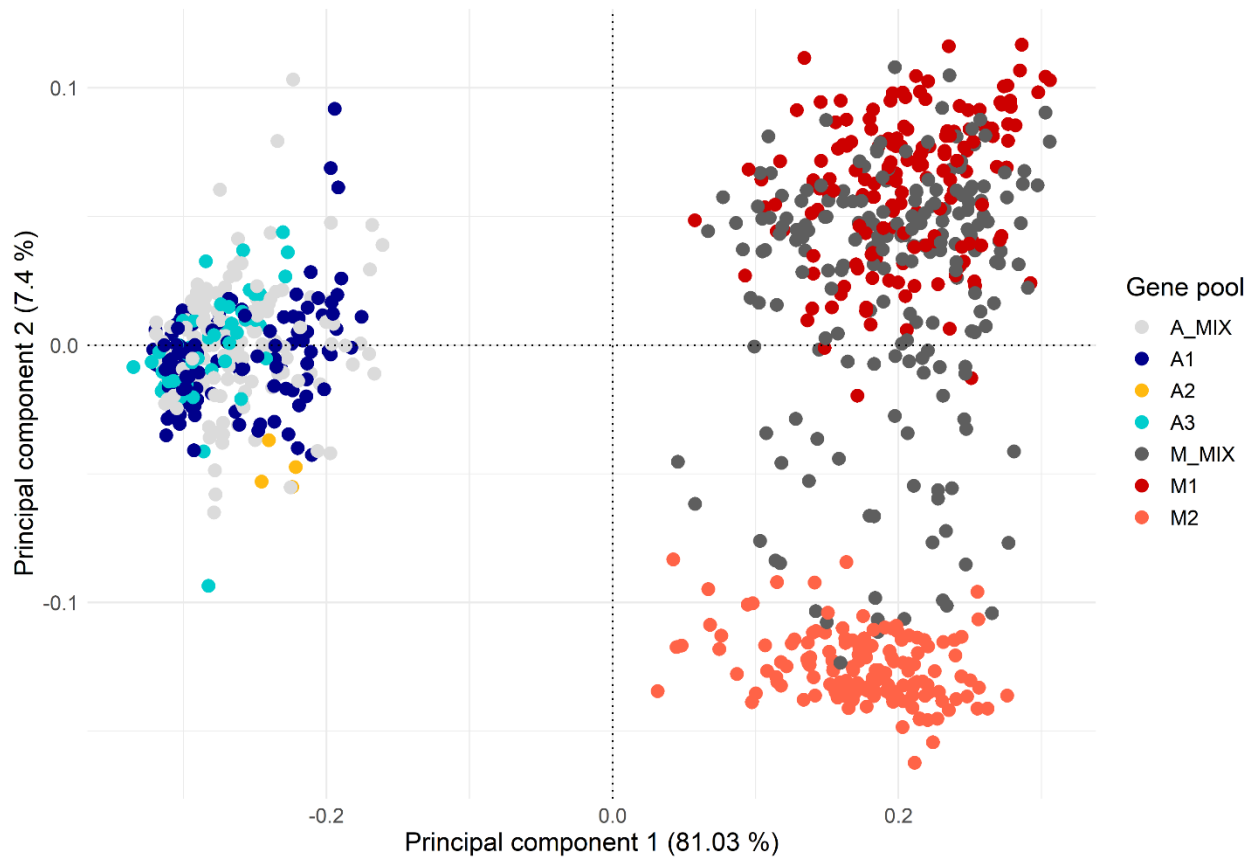
250k thinning

1,701 SNPs

Genotypic analysis

Plant material gene-pool composition

PCA of Mesoamerican and Andean samples





From 1,126 lines + 1 common control variety **thousands** of seed bags are prepared every round (taking into account specific request of citizens for cultivation in field, yards, orchards, gardens, greenhouses, or balcony, terraces, home, indoor...)
Each line is replicated from 10 to 40 times

Before shipping two steps are performed:

- ✓ **Randomization** of CSE lines to prepare the unique postal package for each citizen
- ✓ **Grouping** for shipments to EU, in order to distribute the randomization blocks to the different EU geo-climatic areas

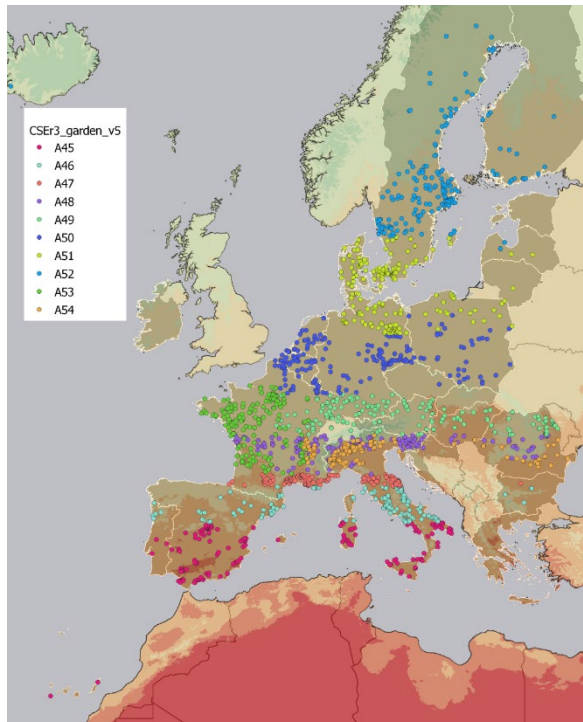


Randomization

One Randomization Unit of CSE lines is based on a set of 1,000 lines; each unit of 1,000 different lines (RU) is independently randomized to define **randomization blocks of 200 combinations of 5 varieties** (for 200 citizens)

In each 1,000 lines randomization (RU), a given line is combined with any other line only once as maximum (*incomplete block design*)

	ROUND 2 new randomization blocks or unit	From bag	To bag	total bags in the RU	Type of bag
1	A25	3501	3700	200	10 seeds
2	A26	3701	3900	200	10 seeds
3	A27	3901	4100	200	10 seeds
4	A28	4101	4300	200	10 seeds
5	A29	4301	4500	200	10 seeds
6	A30	4501	4700	200	10 seeds
7	A31	4701	4900	200	10 seeds
8	A32	4901	5100	200	10 seeds
9	A33	5101	5300	200	10 seeds
10	A34	5301	5500	200	10 seeds
11	A35	5501	5700	200	10 seeds
12	A36	5701	5900	200	5 seeds



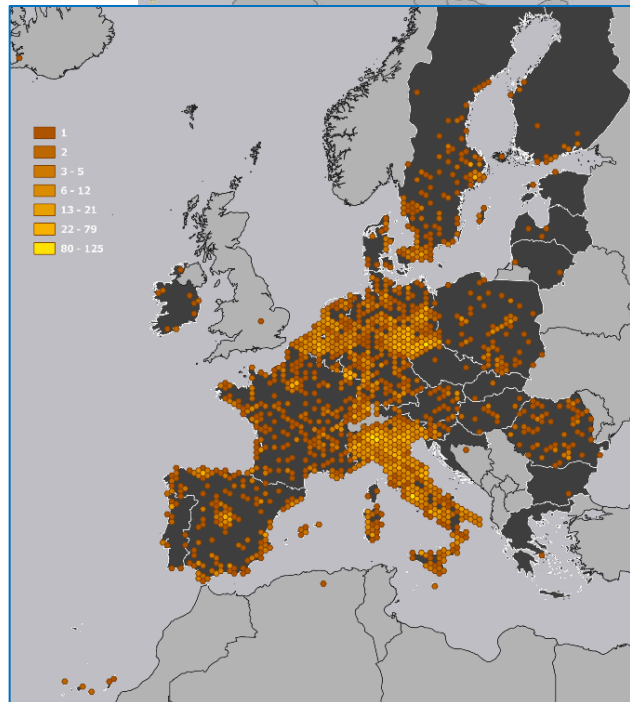
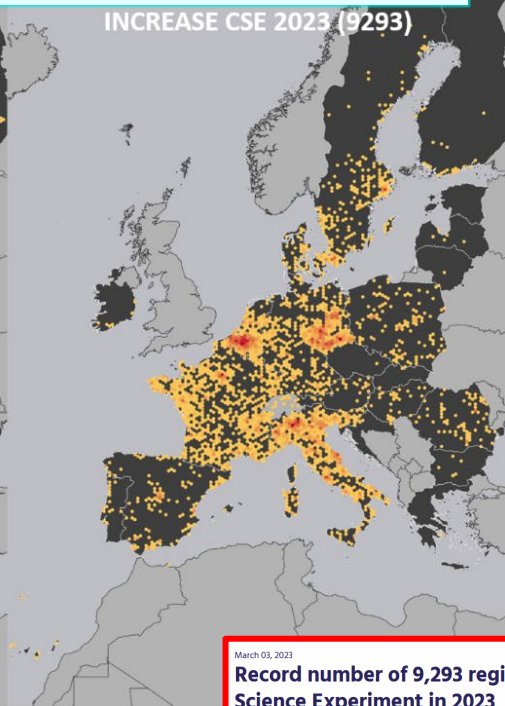
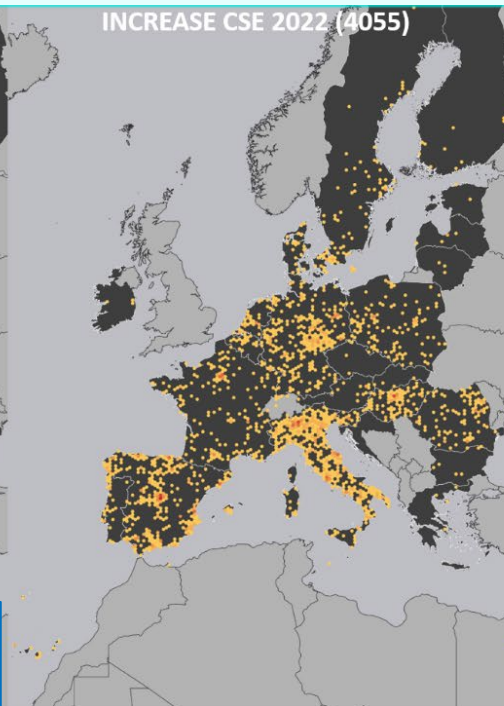
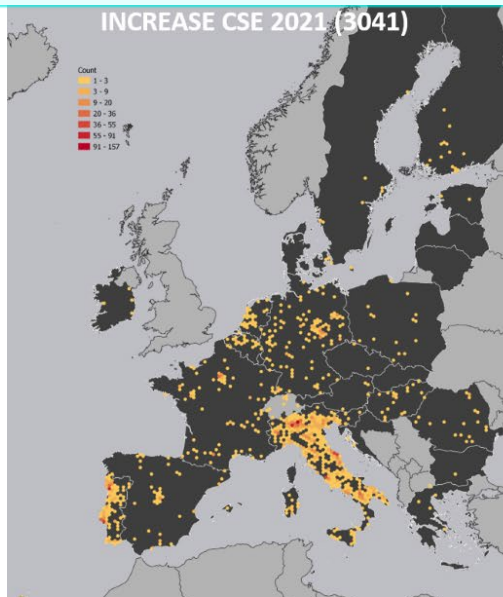
Grouping

All the different citizens sites (geo coordinates) are used to fill the RU

→ Main criteria to group citizen and fill the RUs:

- **Latitude** (day length important for flowering and transition from vegetative to productive phase) and
- **Temperatures and Precipitations** (during common bean growing season)

Common Bean Citizen Science – The registered Citizens



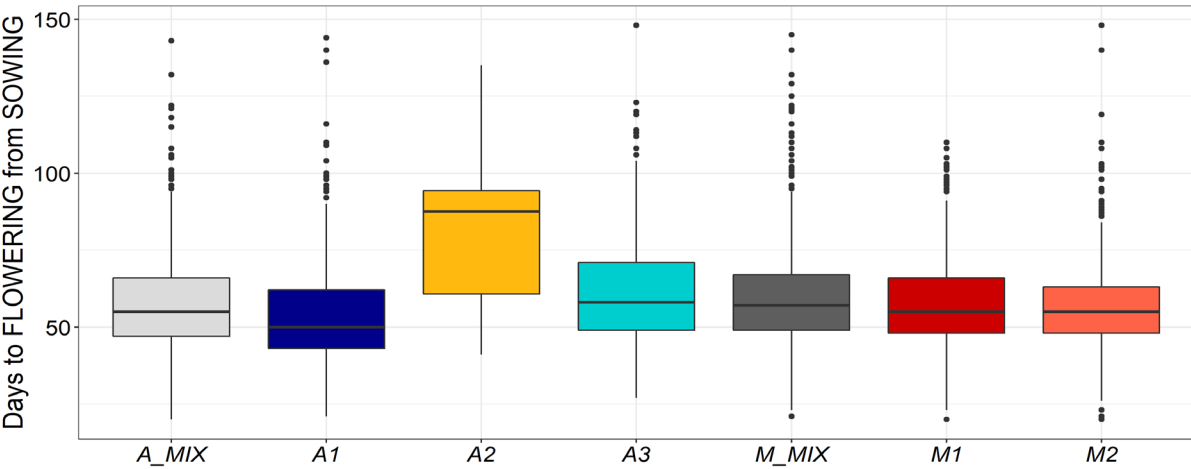
Images by Markus Oppermann P2 - IPK

March 03, 2023
Record number of 9,293 registrations for Citizen Science Experiment in 2023
Registrations for the 3rd round of the INCREASE Citizen Science Experiment are now closed and the bean growing journey for 2023 is about to start with a record number of 9,293 participants from all over Europe. This is about 2,000 more participants than in the first two rounds of the experiment, in 2021 and 2022, combined.

INCREASE CSE 2024
5000 NEW + over 1300
continuing participants

Distribution of days from sowing to FLOWERING of the different Gene pools

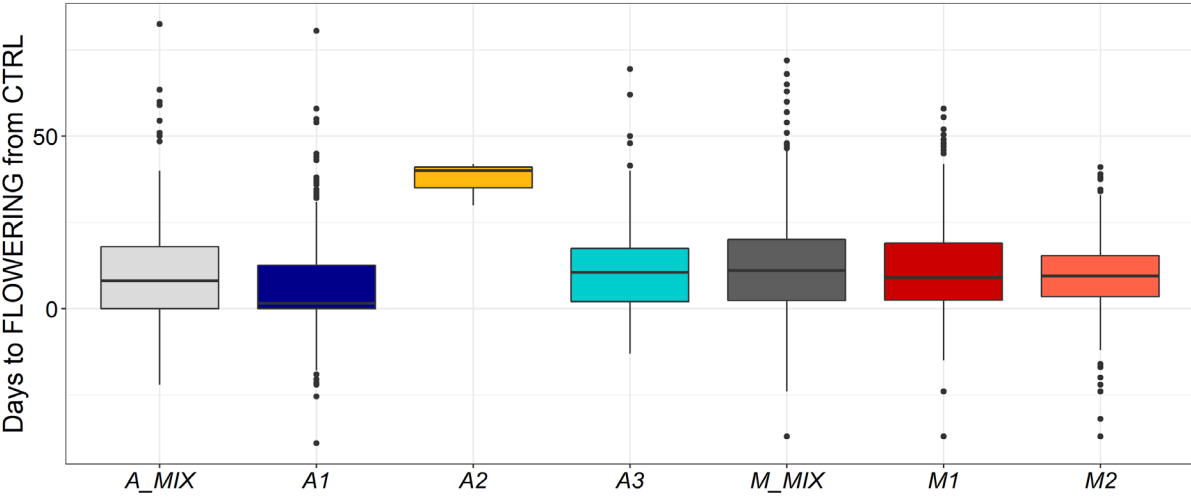
ROUND 3



Number of datapoint for each gene pool 0.8 membership threshold	
■	A1 - 1,040
■	A2 - 12
■	A3 - 310
■	And mix - 1,021
■	M1 - 840
■	M2 - 1,028
■	Meso mix - 1,291

Distribution of days from sowing to FLOWERING from CTRL of the different Gene pools

ROUND 3

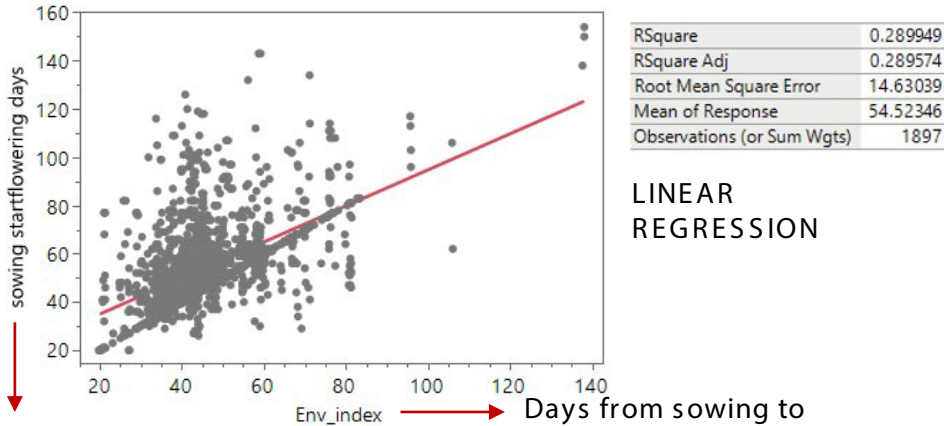


Number of datapoint for each gene pool 0.8 membership threshold	
■	A1 - 400
■	A2 - 3
■	A3 - 111
■	And mix - 374
■	M1 - 321
■	M2 - 395
■	Meso mix - 475



Flowering GWAS – explorative results with Round2 data

1) Evaluate the environmental effect



LINEAR REGRESSION

Days from sowing to flowering of other genotypes grown by the same citizen

Days from sowing to flowering of INCBN_00000

- Each Citizen was considered as one environment
- Overall 402 Citizens had data for INCBN_00000
- Environmental index was created sorting citizens according to days to flowering of the CTRL (INCBN_00000)

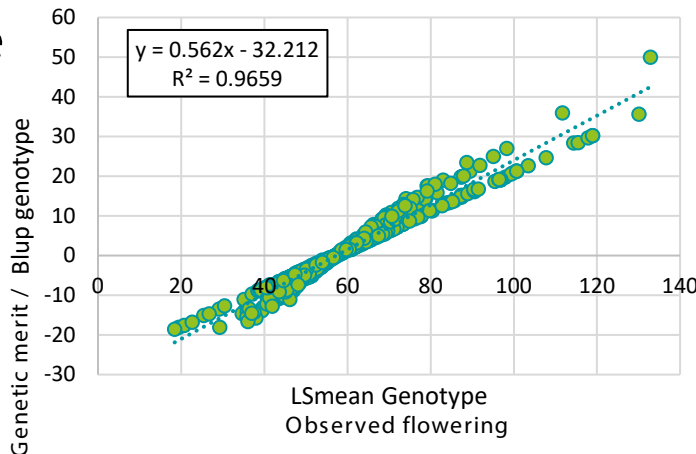
669 genotypes

- Grown by the same citizen together with CTRL
- GBS data available

2) Heritability estimation of Flowering

$$Y = G + E_{ind}^* + e$$

↓ ↓
Rand - BLUP Fixed



Random effect

Term	% tot variance	H ²
Days After Sowing	49.0	0.56

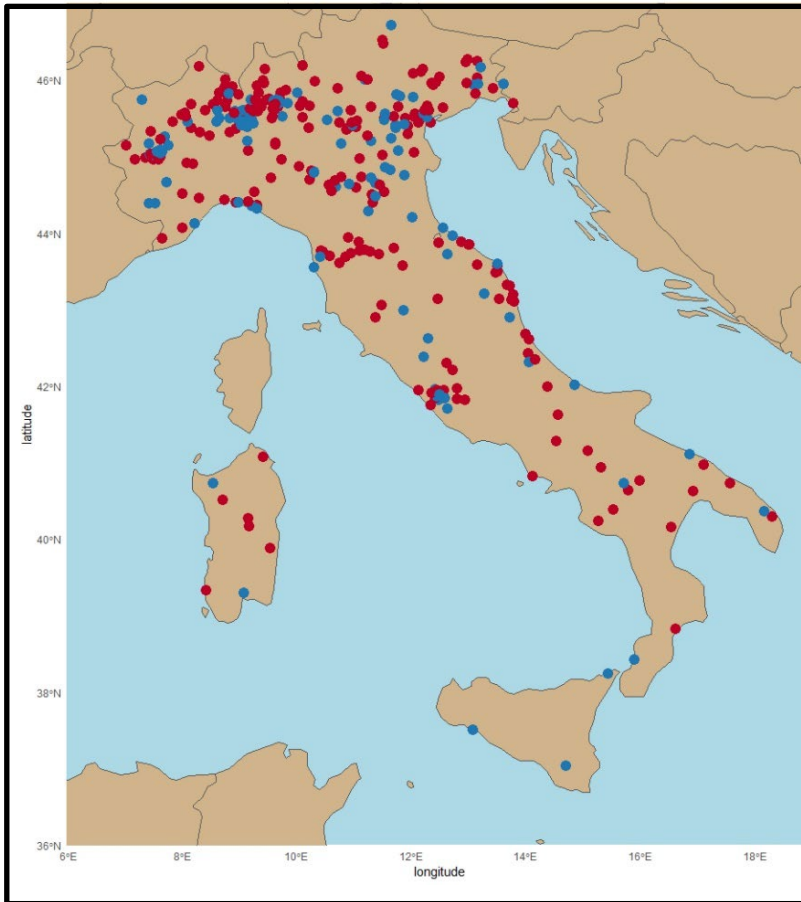
Fixed effect

Term	Significance
E _{ind}	**

Heritability was measured as the regression coefficient between BLUP (genetic merit) and phenotypic trait (observed flowering)

Selection of the coordinates for the extraction of the environmental variables

336 geographic coordinates (each linked to a Citizen) have been selected



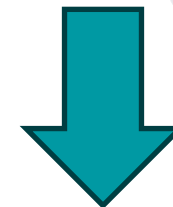
Coordinates selected based on citizens, divided according to whether the experiment was conducted in the field (red) or in a pot (blue)

Citizens who participated in the second and/or third rounds of the experiment in Italy

Both field and pot data were included

Those who entered at least one data for flowering and harvest

Those who had reported the data for the control line INCBN_00000



Only one coordinate was selected when two or more Citizens were located within less than 1 km

Geographical Coordinates and Citizens location



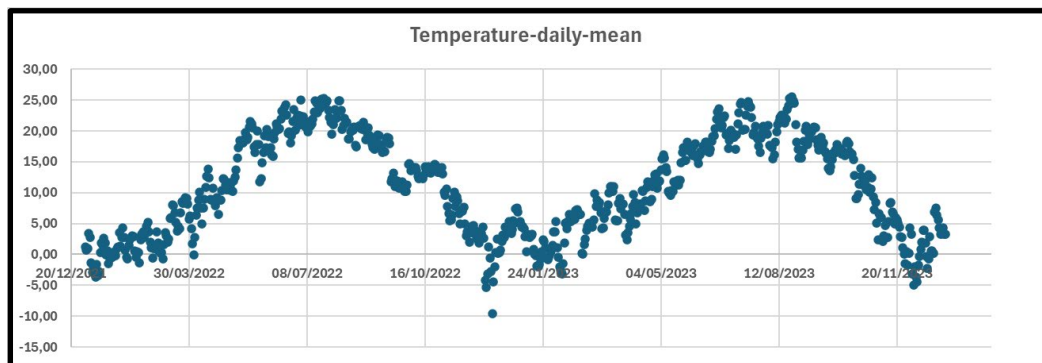
In addition to the 336 Citizens/location that we used as a reference to obtain environmental data, additional 140 citizens were located within a distance of less than 1 kilometer from the selected coordinates (Citizens) (red spot)

Environmental variables (1 km resolution)

	A	B	C
1	index	time	temperature-daily-mean
2	0	01/01/2022	1,22
3	1	02/01/2022	0,72
4	2	03/01/2022	0,99
5	3	04/01/2022	3,38
6	4	05/01/2022	2,78
7	5	06/01/2022	-1,40
8	6	07/01/2022	-2,92
9	7	08/01/2022	-1,99
10	8	09/01/2022	-1,69
11	9	10/01/2022	-3,73
12	10	11/01/2022	-1,68
13	11	12/01/2022	-3,37
14	12	13/01/2022	-2,74
15	13	14/01/2022	0,36
16	14	15/01/2022	1,85
17	15	16/01/2022	1,47
18	16	17/01/2022	2,65
19	17	18/01/2022	1,83
20	18	19/01/2022	-0,11
21	19	20/01/2022	0,59
22	20	21/01/2022	-1,52
23	21	22/01/2022	-0,75
24	22	23/01/2022	-0,29
25	23	24/01/2022	-0,41

Environmental Variables for each coordinate:

- Temperature daily mean: °C
- Temperature daily maximum: °C
- Temperature daily minimum: °C
- Precipitation: mm/day
- Incoming solar radiation: Watts (W/m²)
- Soil texture (relative composition of sand, silt, and clay and soil classification)

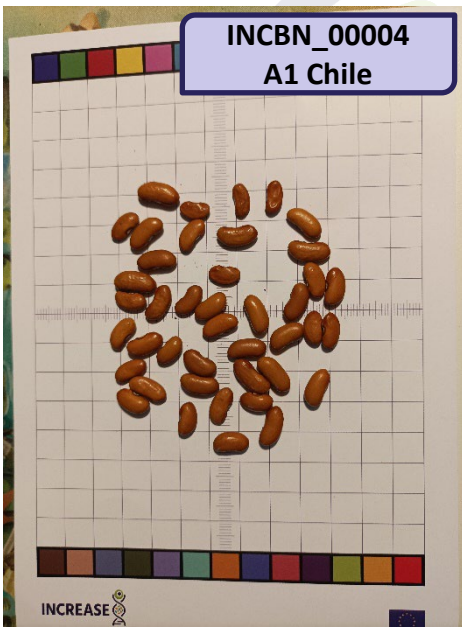


For each location a daily timeseries for each environmental variable, from 01/2022 to 12/2023





TOP 50%
most sown lines

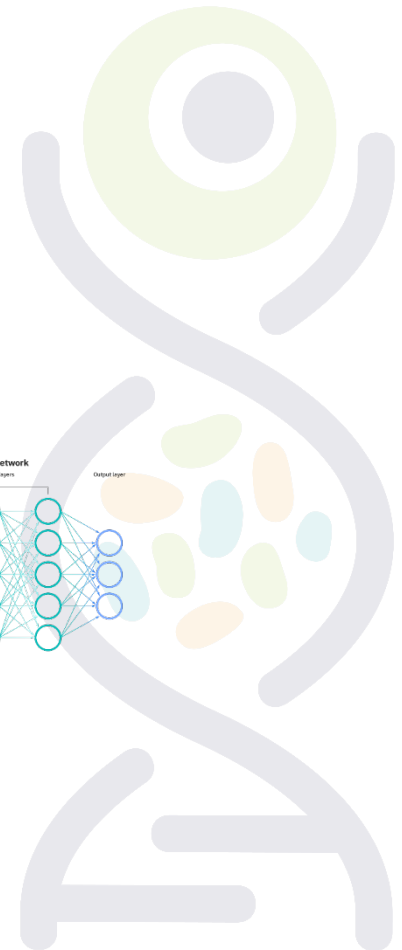
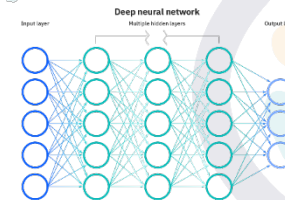
INCREASE project code	Country	Race	Climb or Bush	N. TIMES SOWING	N. TIMES FLOWERING	N. TIMES HARVEST	Number of seed exchanges	RATIO N.EXCH/N.SOWING
INCBN_01804	Georgia	M1	climbing	16	5	5	9	0.56
INCBN_03427	NA	MESO_MIX	climbing	16	4	10	9	0.56
INCBN_02167	Georgia	M1	climbing	15	5	2	6	0.40
INCBN_02226	Georgia	MESO_MIX	climbing	15	9	5	5	0.33
INCBN_00076	Costa Rica	M2	climbing	40	13	11	10	0.25
INCBN_00177	Italy	A3	climbing	17	9	6	4	0.24
INCBN_00004	Chile	A1	bush	44	23	21	10	0.23
INCBN_00232	Mexico	AND_MIX	bush	42	13	18	9	0.21
INCBN_06963	Yugoslavia	M1	climbing	19	3	3	4	0.21
INCBN_10212	NA	A3	bush	15	4	2	3	0.20
INCBN_00698	Georgia	<i>in progress</i>	climbing	15	5	7	3	0.20
INCBN_07165	Spain	MESO_MIX	climbing	15	8	7	3	0.20
INCBN_06493	Ethiopia	<i>in progress</i>	bush	42	22	18	8	0.19
INCBN_01524	Cuba	M2	climbing	21	6	7	4	0.19
INCBN_01610	Cuba	M2	bush	53	22	17	10	0.19
INCBN_01637	Cuba	M2	climbing	54	25	18	10	0.19
INCBN_00199	Germany	AND_MIX	bush	55	22	13	10	0.18
INCBN_00468	Italy	MESO_MIX	climbing	44	17	10	7	0.16
INCBN_00139	France	A1	bush	52	24	23	8	0.15
INCBN_00231	Guatemala	A1	bush	40	12	11	6	0.15
INCBN_01512	Cuba	M2	bush	34	17	15	5	0.15
INCBN_02098	Georgia	MESO_MIX	climbing	52	19	14	7	0.13
INCBN_00257	Dominican Republic	A1	bush	15	5	5	2	0.13
INCBN_00967	Bulgaria	<i>in progress</i>	climbing	15	7	6	2	0.13
INCBN_06452	Cuba	AND_MIX	climbing	15	2	2	2	0.13
INCBN_06455	Cuba	MESO_MIX	climbing	15	4	5	2	0.13
INCBN_03278	Slovakia	MESO_MIX	climbing	15	15	11	2	0.13
INCBN_00264	Mexico	AND_MIX	bush	53	20	15	7	0.13
INCBN_00473	Albania	AND_MESO_MIX	climbing	46	10	9	6	0.13
INCBN_00055	Mexico	MESO_MIX	climbing	46	19	14	6	0.13
INCBN_00336	Peru	AND_MIX	bush	16	6	2	2	0.13
INCBN_01496	Cuba	M2	climbing	16	8	6	2	0.13
INCBN_01741	Georgia	M1	climbing	16	0	0	2	0.13
INCBN_00061	Mexico	M2	climbing	48	11	9	6	0.13
INCBN_00348	NA	MESO_MIX	climbing	16	7	6	2	0.13

TOP 5%
ratio
sown/exchanged



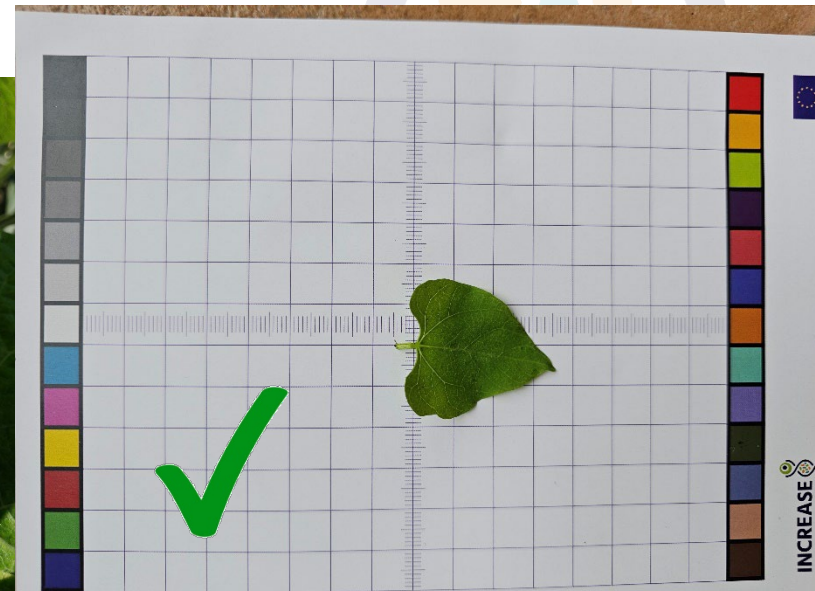
Labelling and training of a model to infer phenotypes from images: How to automate the processing?

- Grab data 
- Take some expert 
- Label high quality Images (this takes time) 
- Train a deep learning model (this takes time)
- Play with it 



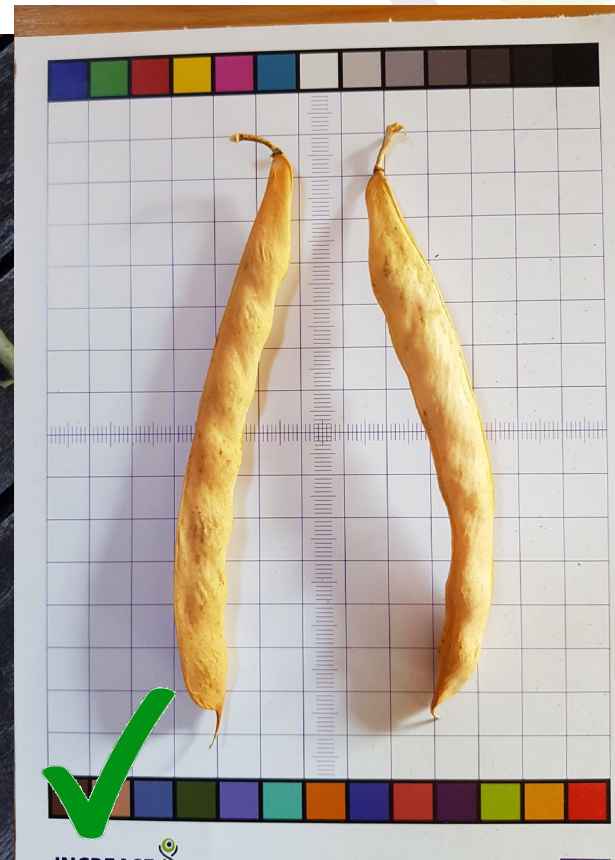
Use a color checker on the background

- Where possible try to use the INCREASE color-checker.
- The use of color checker allows use to calibrate the colors and make measurements.
- Please avoid shadows where possible
- Please try to keep hands out of the color checker...we know this is sometimes complex ;-)
- Be sure that color checker is fully contained in the image



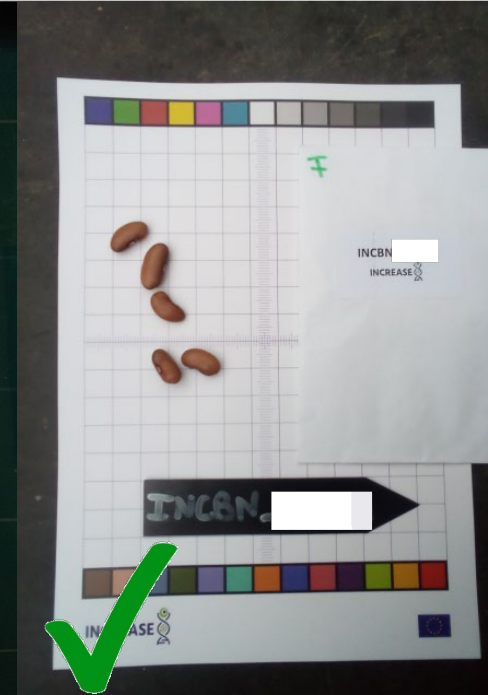
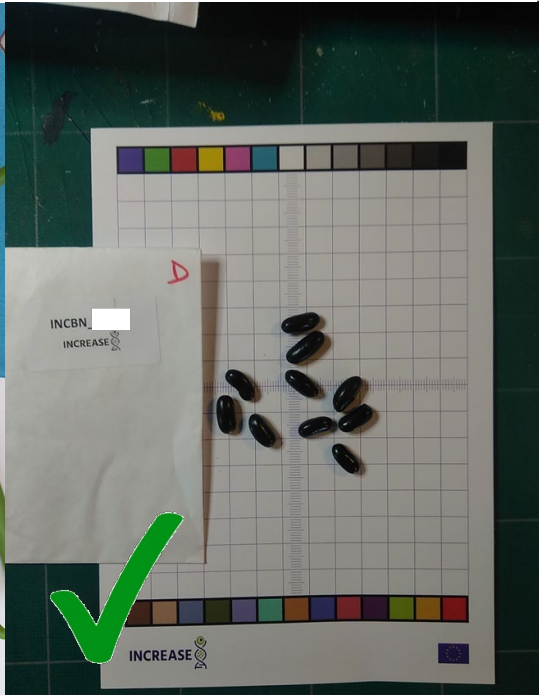
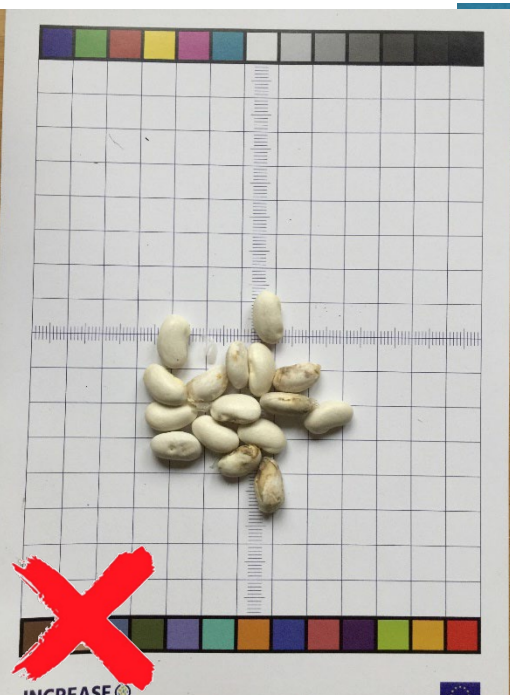
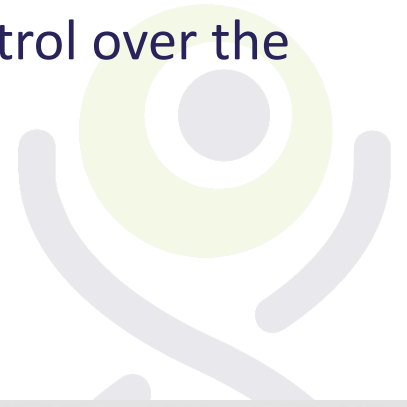
Too many «objects»

- Please, try to consider only few objects....as you see we can not see some color and the calibration grid...



Items stuck together

- Try to arrange objects with «care»...
- Try to individually place and space the beans / pods. This aspect is of-course time-consuming but allows for complete control over the composition and spacing of the beans/pods.



Next...

AI to define seed-related traits, and Seeds validation

- Set up of a model for seed traits definition
- Seed images labelling and training of a model that given an image it allows to assign an identity % to an accession



photo documentation of an SSD cycle 1 bean harvest

**Seeds
validation**




Seeds validated by a Citizen

INCREASE

Intelligent Collections of Food Legumes Genetic Resources for
European Agrofood Systems



***Thanks for your
attention!***



Roberto Papa

