Mega-environment analysis & Breeding for ME-specific oat cultivars for Canada

Weikai Yan

Ottawa Research and Development Center

Agriculture and Agri-Food Canada

March 2022

Why mega-environment analysis?

- To exploit specific adaptation
- To improve selection reliability
- To enhance breeding progress
- To increase regional and global production

What are mega-environments?

- Subregions of a crop species' growing regions

Subregions... defined by geographical or environmental factors

- By latitude: Tropical, temperate, polar regions
- By longitude: eastern vs. western Canada
- By flowering time of three plant species
  - 3 Zones in Quebec (Dube´ and Chevrette, 1978)
- By corn heat unit
  - 6 CHU areas of Ontario (Major et al., 1983)
- By soil type
  - 4 soil zones in Canadian prairies

Widely used but may not be accurate enough to guide breeding and cultivar recommendation

Subregions… based on data from multi-environment trials

The criterion is "which-won-where"

- "If the winning cultivar is different in different environments, then the environments should be divided into different mega-environments"

Gauch & Zobel, crop Sci. 1997

The basis is "GGE"
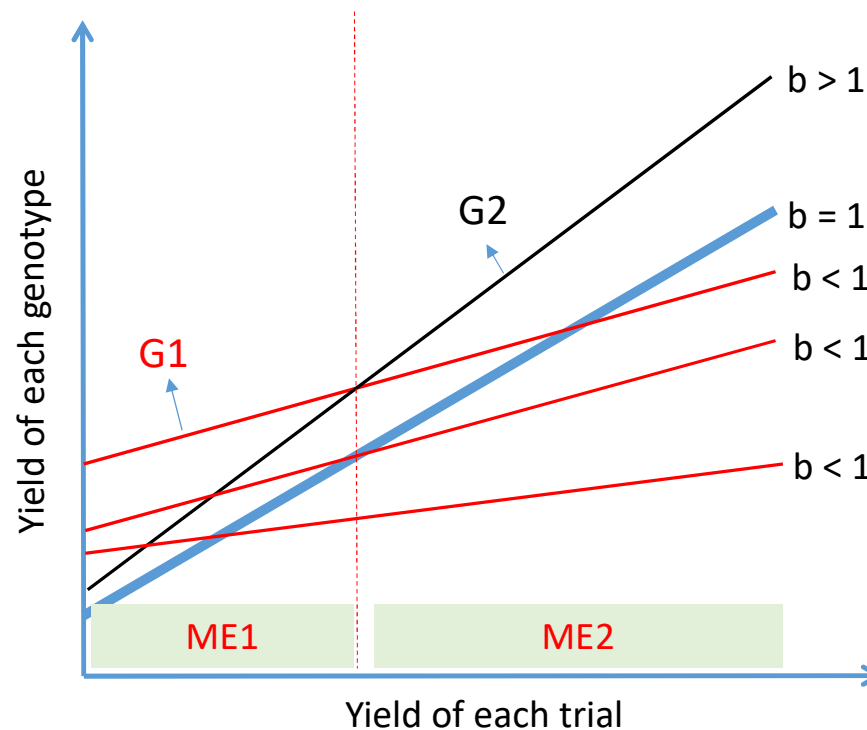
- P = E + G + GE

- G and GE are relevant and must be considered for
  - Genotype evaluation,
  - Environment evaluation
  - Mega-environment analysis

- "GGE" means…
  - G + GE
  - G/GE
  - G vs GE

(Yates and Cochran, 1938; Yan et al. 2000)

Graphical displays of "GGE" and "which-won-where"

- The joint regression graph
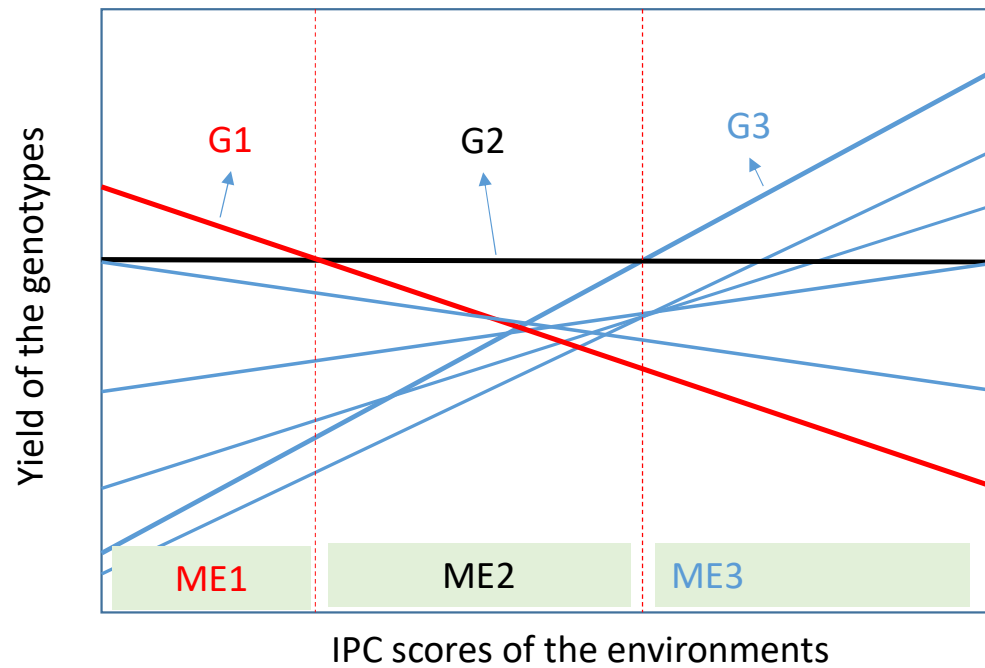- The AMMI graph
- The GGE biplot

Regression on means of the environments to show which-won-where

Each genotype is displayed as a linear regression against the means of the environments (trials)

- Parallel: no GE
- No crossover: minor GE
- Crossover GE: rank change
- "Which-won-where": meaningful crossover GE, which defines ME

(Yates & Cochran, 1938; Finlay & Wilkinson, 1963; Eberhart & Russell, 1966; Ceccarelli, 1989)
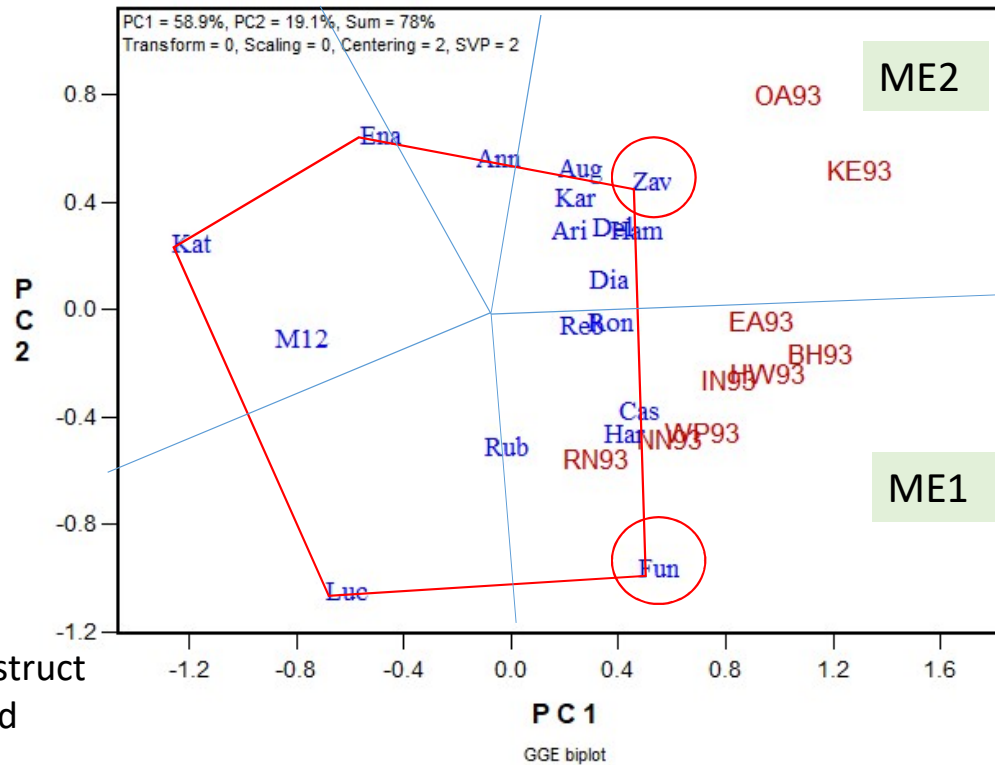
Regression on the IPC1 scores of the environments (AMMI1 graph) to show which-won-where

Each genotype is displayed as a linear regression against the IPC1 scores of the environments

IPC1: the 1st principal component of GE

Gauch and Zobel, 1997, Crop Sci.

GGE biplot to show which-won-where

- Explains more G+GE
- More straightforward to construct
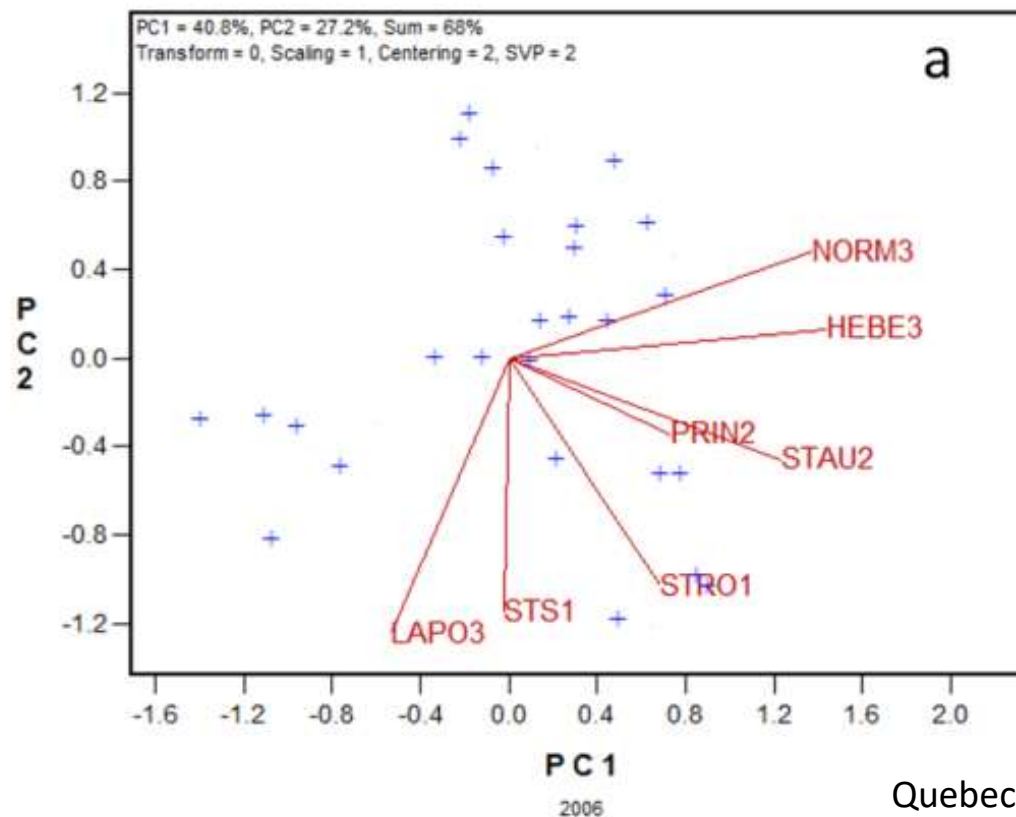- Allows more entries displayed
- More versatile functions
- More elegant

Biplot: Gabriel, 1971; GGE biplot: Yan, 2001

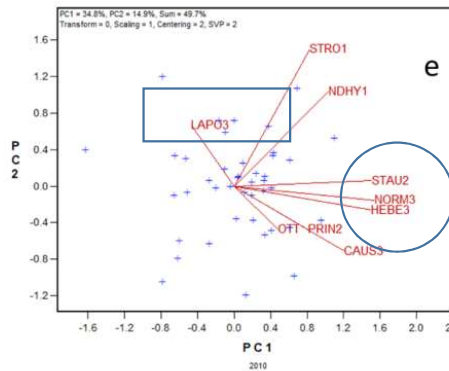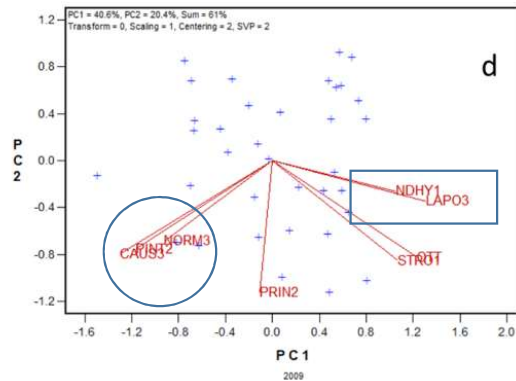# Mega-environment delineation must be based on <u>repeatable patterns</u>
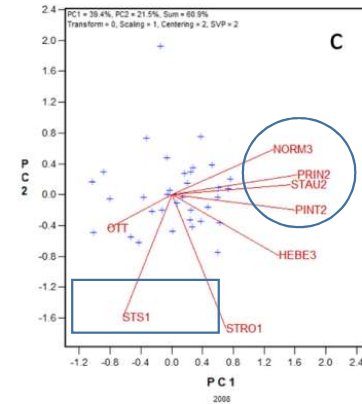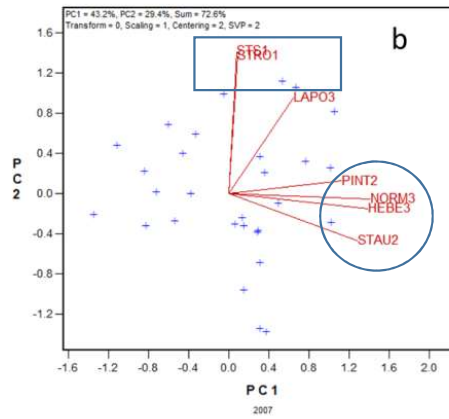
- How to assess the repeatability?
  - "Analyze yearly and summarize across years" (DeLacy et al., 1996)
  - Easy said, hard to do

- GGE+GGL biplot (Yan, 2014, 2015)
- LG biplot (Yan, 2019)

GGE biplot to display the genetic correlations between locations in a year

- r = cosine of the angle between two environments
  - r = 1 (GE = 0)
  - r = 0 (GE = G)
  - r = -1 (G = 0)

  (different alleles of opposite functions)

PC1 = 40.8%, PC2 = 27.2%, Sum = 68%
Transform = 0, Scaling = 1, Centering = 2, SVP = 2

a

Quebec 2006

Is the GE pattern repeatable across years?
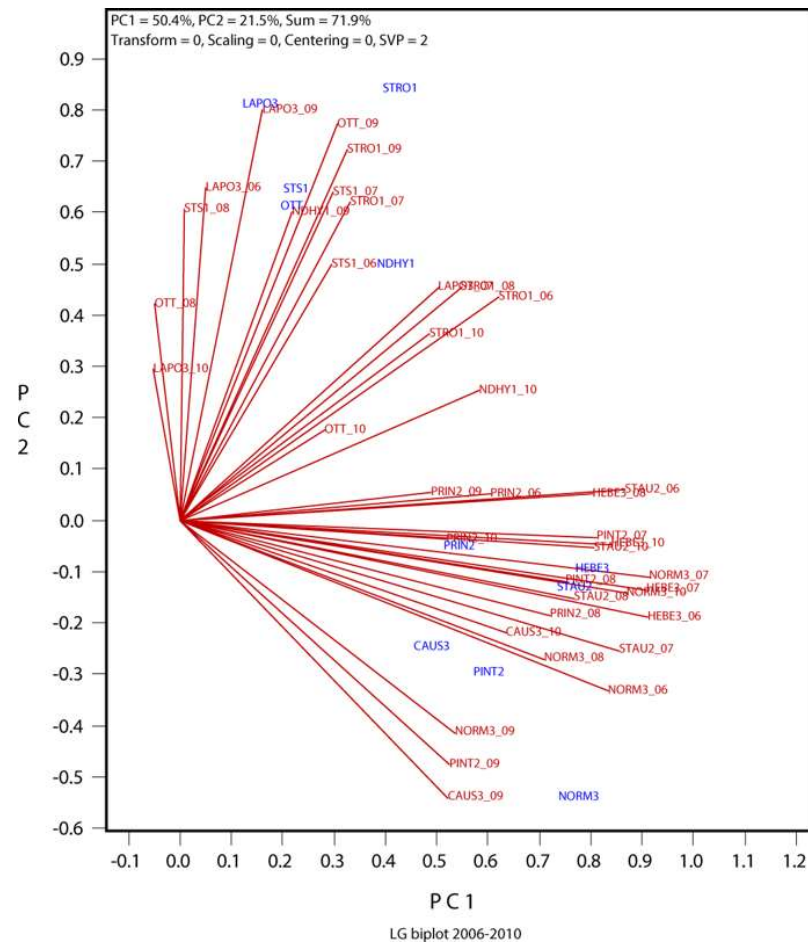

"Analyze yearly and summarize across years"

14
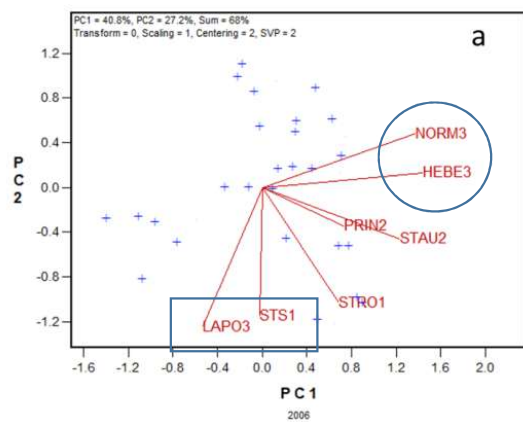
- <u>LG biplot to</u> display the location-relations in multiple years

- LG biplot is a location by trial biplot

- LG: Location Grouping

LG biplot 2006-2010

Yan, 2019, Scientific Reports

Trials at each location are displayed as a cluster...

- Locations that are always positively correlated
  - Same ME
- Locations that are always uncorrelated or negatively correlated
  - Different MEs
- Locations that variably correlated
  - Same ME with large unpredictable GE

Yan, 2019, Scientific Reports
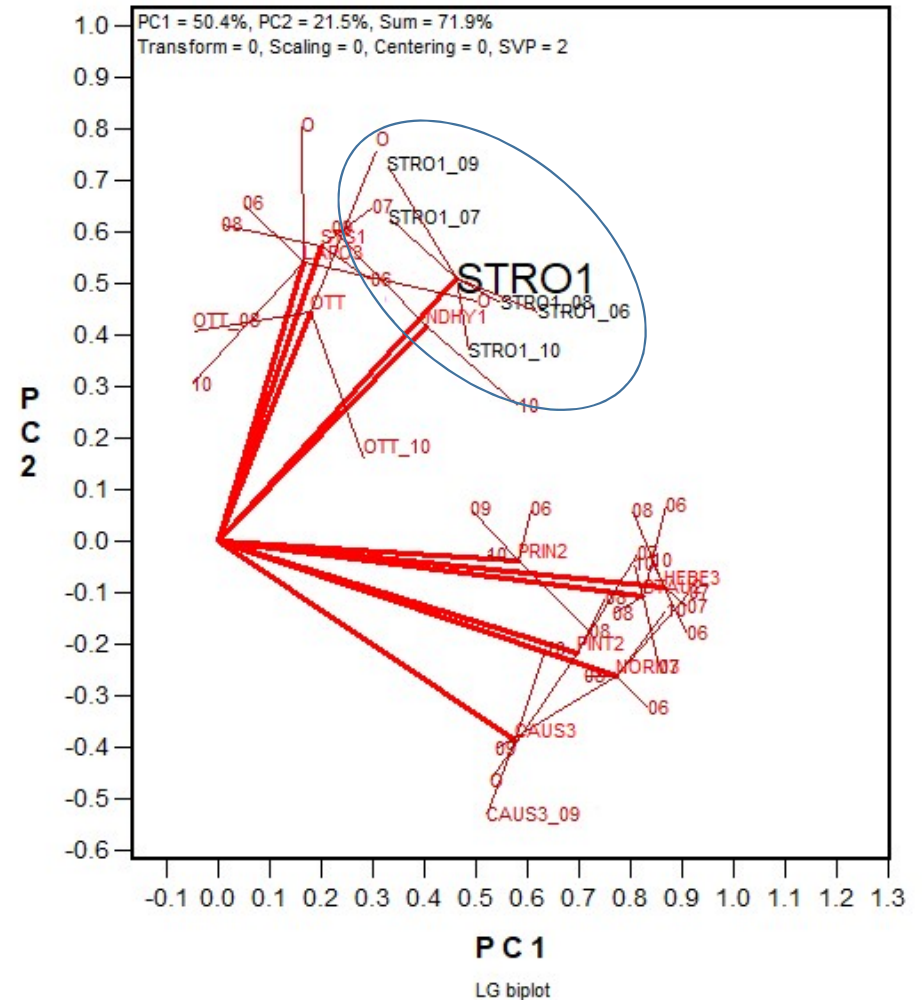
LG biplot

The LG biplot reveals two subregions

- **ME2A**: Zone1 + La Pocatière (Zone3) + OTT
- **ME2B**: Zone2 + Zone3 excluding La Pocatière (Zone3)
- The same pattern year after year (Yan, 2021, Front. Plant Sci.)

Yan, 2019, Scientific Reports
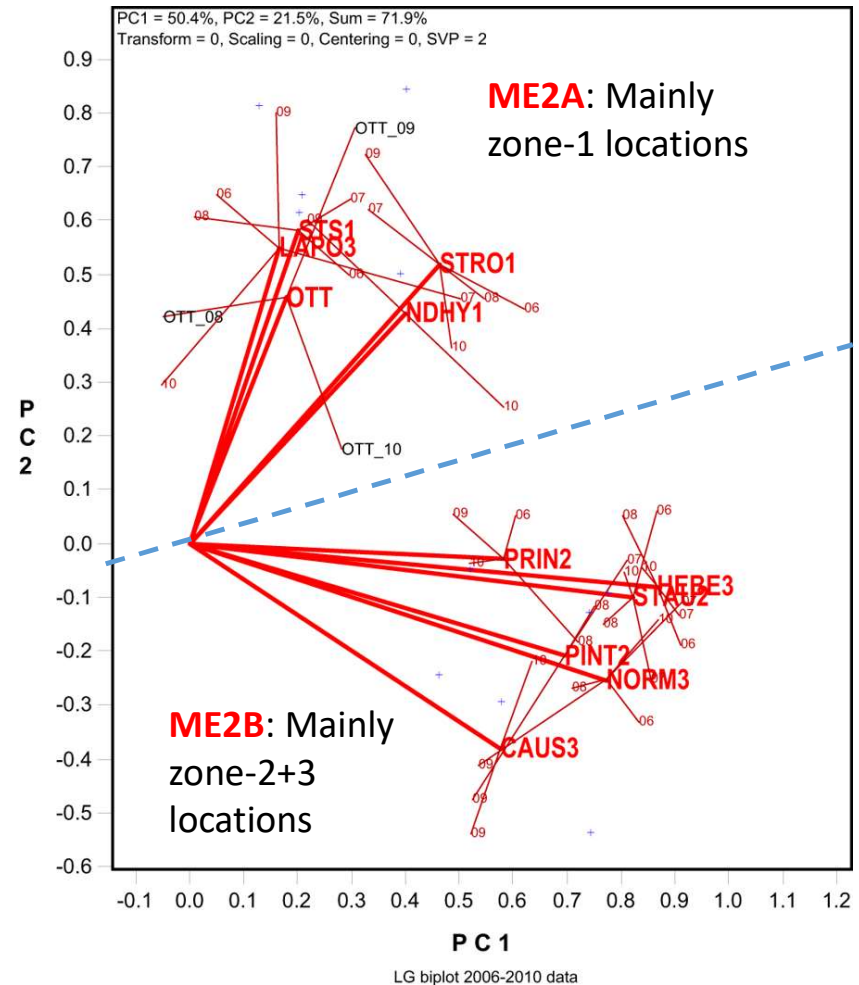
The LG biplot separates <u>repeatable GE</u> from <u>unrepeatable GE</u>

- GE between subregions
  - Repeatable GE = repeatable GL = Genotype by subregion interaction
- GE within subregions
  - Unrepeatable GE = unrepeatable GL + GY + GLY

LG biplot
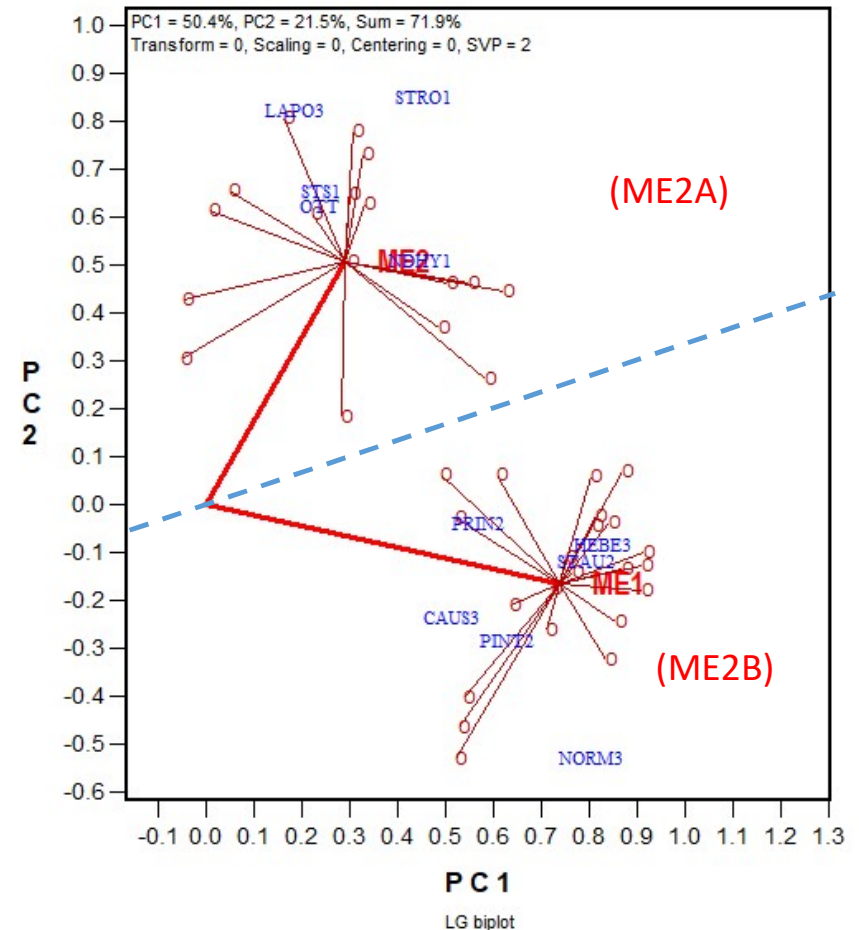
Different mega-environments had different winner cultivars

Quebec Oat yield, 2006-2010
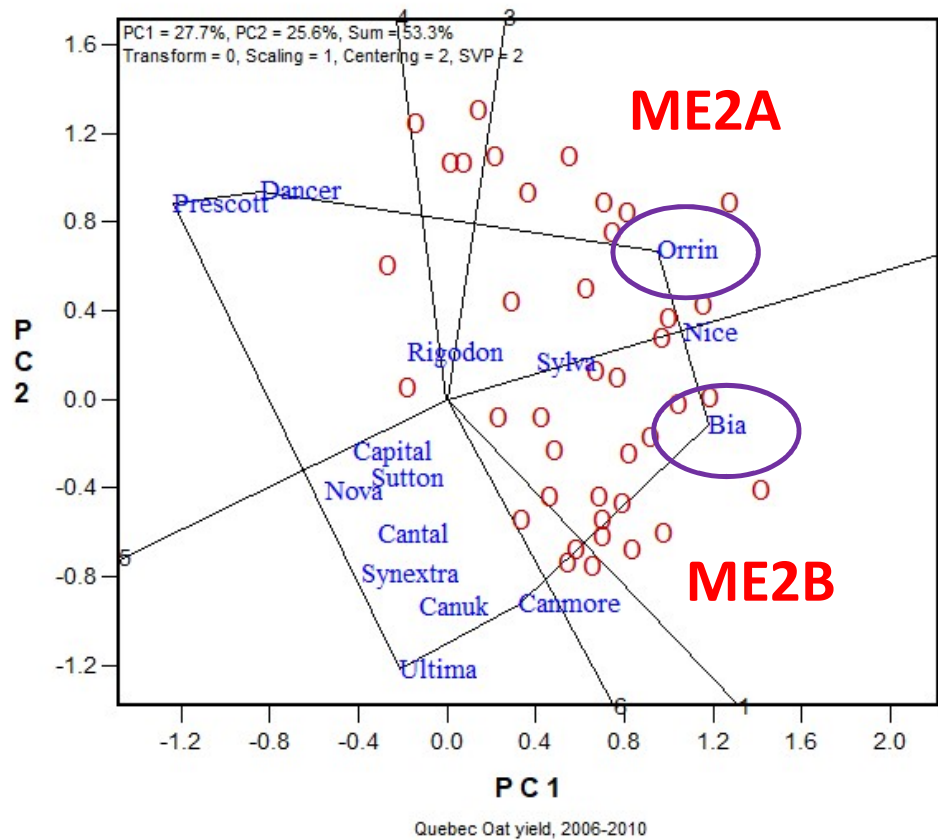
Two contrasting MEs in eastern Canada

- ME1 = Areas 2 & 3 of Ontario

- ME2 = northern Ontario, Quebec, Maritimes

- Best cultivar for ME1: AAC Bullet

- Best cultivar for ME2: AAC Nicolas

Yan et al., 2015, Euphytica
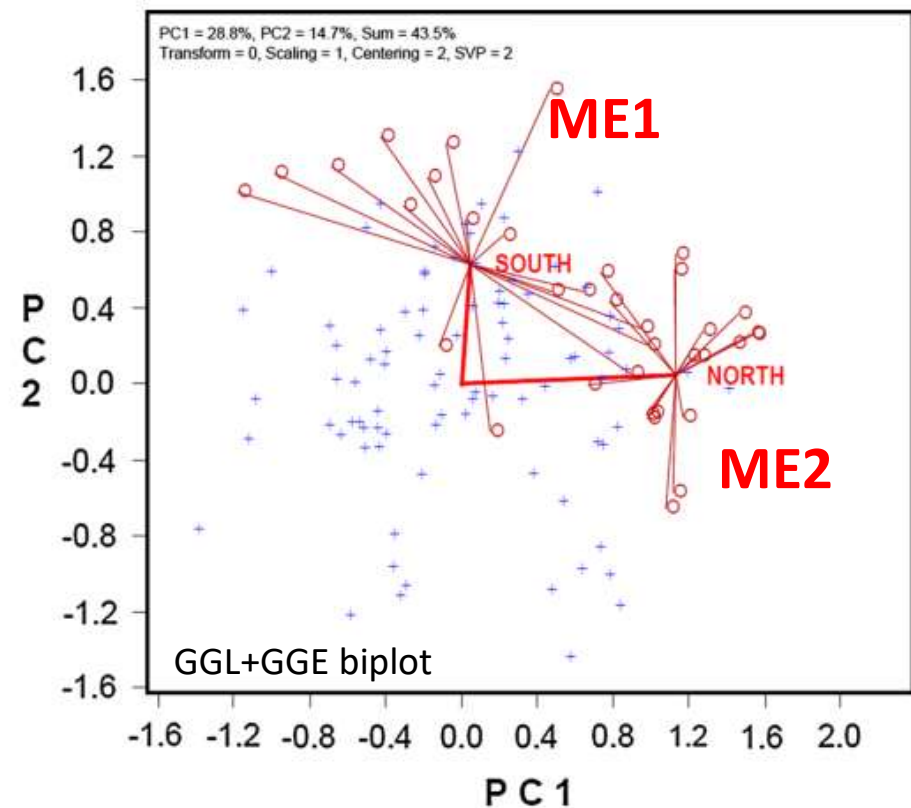
Three oat mega-environments in Canada

ME3

ME2

ME1

Yan et al., 2021, Crop Science
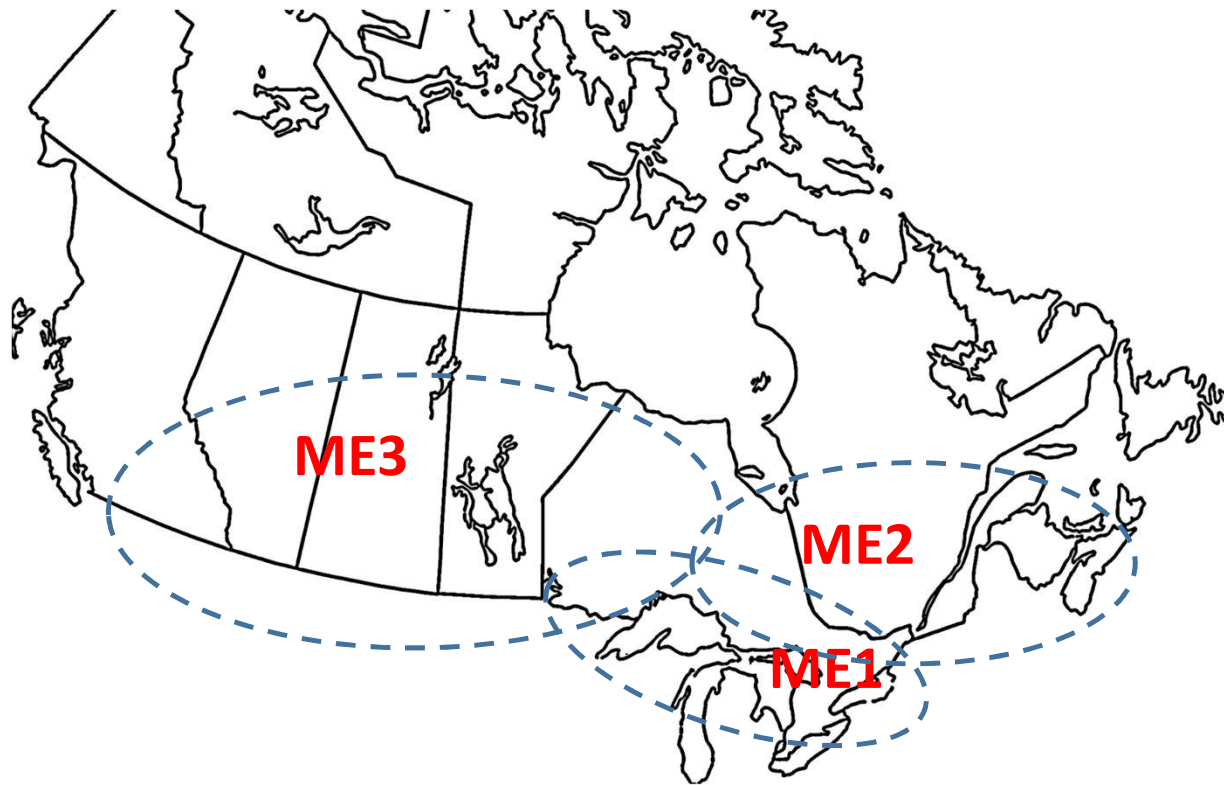
Mega-environments are *not* the same as the agro-ecological zones/areas

- Quebec
  - ME2A = Zone1
  - ME2B = Zone2 + Zone3
  - LAPO3 is more similar to Zone1 rather than Zone3
- Ontario
  - ME1 = Areas 2 & 3
  - ME2 = Areas 5 & 6
- Prairies
  - All 4 soil zones belong to a single ME

Select within mega-environments to improve heritability (reliability of selection)

Variability of the breeding population

$$h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_{GE}^2/ly}$$

Selecting within mega-environments to reduce GE

"Adequate testing" to accommodate unrepeatable GE ($y \leq 3$)

PC1 = 28.8%, PC2 = 14.7%, Sum = 43.5%
Transform = 0, Scaling = 1, Centering = 2, SVP = 2

ME1

ME2

SOUTH

NORTH
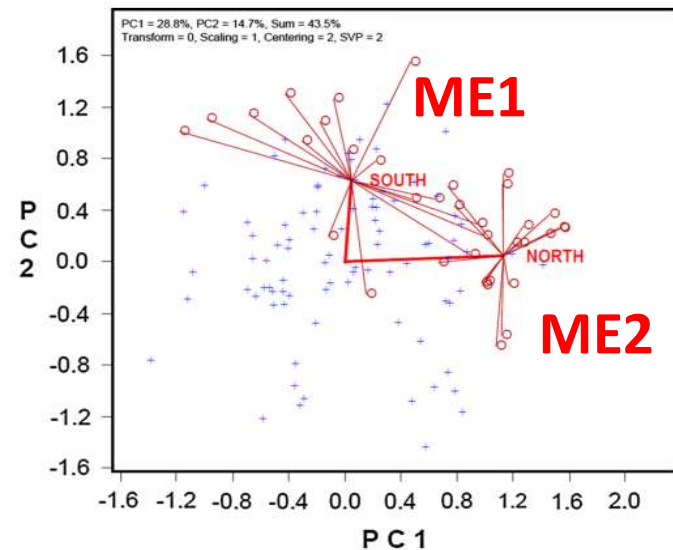
Two aspects in dealing with GE

- Utilizing repeatable GE through mega-environment analysis
  - To reveal repeatable GE and mega-environments
  - To select for each ME
  - Converting repeatable GE to G within ME$\rightarrow$ reduced GE within mega-environments $\rightarrow$ improved heritability $\rightarrow$ improved selection gain$\rightarrow$ maximized local and global production
- Accommodating GE within a mega-environment by "adequate testing"
  - Number of locations
  - Number of years
  - (Number of replicates)

Dealing with GE through Genomic Selection

Variability of the breeding population

$$h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_{GM}^2/m}$$

Use of ME-specific GS models to reduce genotype by model interaction (each model represents a training dataset)

Number of GS models (or training datasets)
$(m \rightarrow \quad)$

Yan, 2021, Front. Plant Sci.

Heritability in the Breeder's Equation

Selection intensity, should be decided by the population size

Sq. root of Heritability

$$B = (\mu + i h \sigma_G)/Y$$

Mean and variability: the usefulness of the breeding population

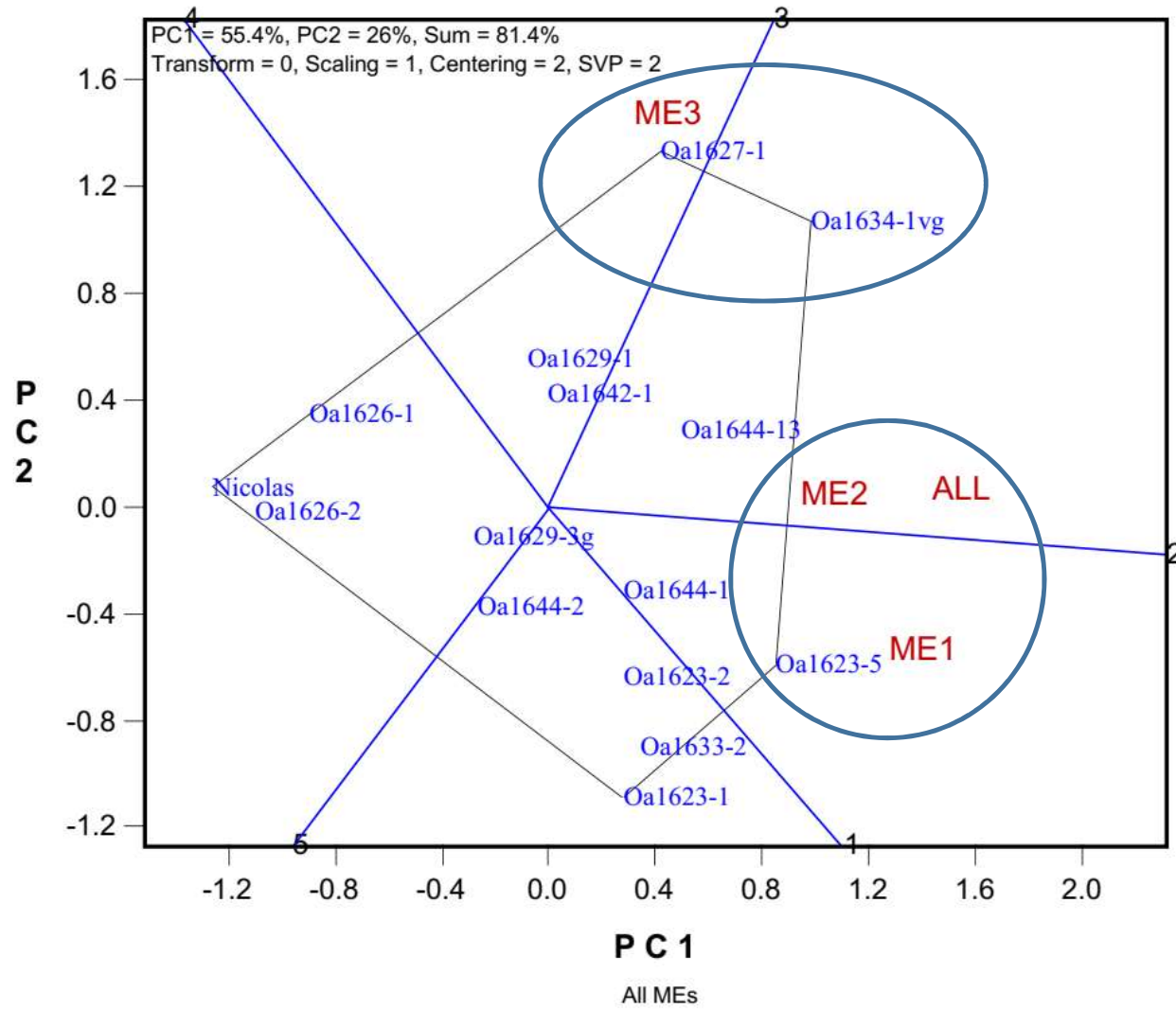The Breeder's equation (Eberhard, 1970; Yan, 2021)

Steps to breed ME-specific oat cultivars for Canada

- Understanding the target region through mega-environment analysis
- Developing a hopeful breeding population
  - Crosses between breeding lines/cultivars adapted to different MEs
- Exploratory trials covering all MEs to identify broadly and/or specifically adapted genotypes
  - 3 to 5 Locations in Areas 2 and 3 of Ontario (ME1)
  - 3 to 4 Locations in the other regions of eastern Canada (ME2)
  - 2 to 3 Locations in the Canadian prairies (ME3)
- Registration trials targeting specific MEs
  - Ottawa RDC registration trials for ME1 (and ME2)
  - Quebec registration trials for ME2
  - Western Cooperative Oat Registration Trials (WCORT) for ME3

A real story...

- Crosses made in 2013
- 66 breeding lines were tested across Canada in 2018
- 17 lines were further tested across Canada in 2019

2018 Preliminary (10)
2019 Registration (10)

PC1 = 55.4%, PC2 = 26%, Sum = 81.4%
Transform = 0, Scaling = 1, Centering = 2, SVP = 2

ME3
Oa1627-1
Oa1634-1vg
Oa1629-1
Oa1642-1
Oa1626-1
Oa1644-13
Nicolas
Oa1626-2
ME2   ALL
Oa1629-3g
Oa1644-1
Oa1644-2
ME1
Oa1623-5
Oa1623-2
Oa1633-2
Oa1623-1

PC1

All MEs

2018 Preliminary (10)
2019 Registration (10)

PC1 = 72.4%, PC2 = 24.8%, Sum = 97.2%
Transform = 0, Scaling = 1, Centering = 2, SVP = 2

ME1 vs. ME2
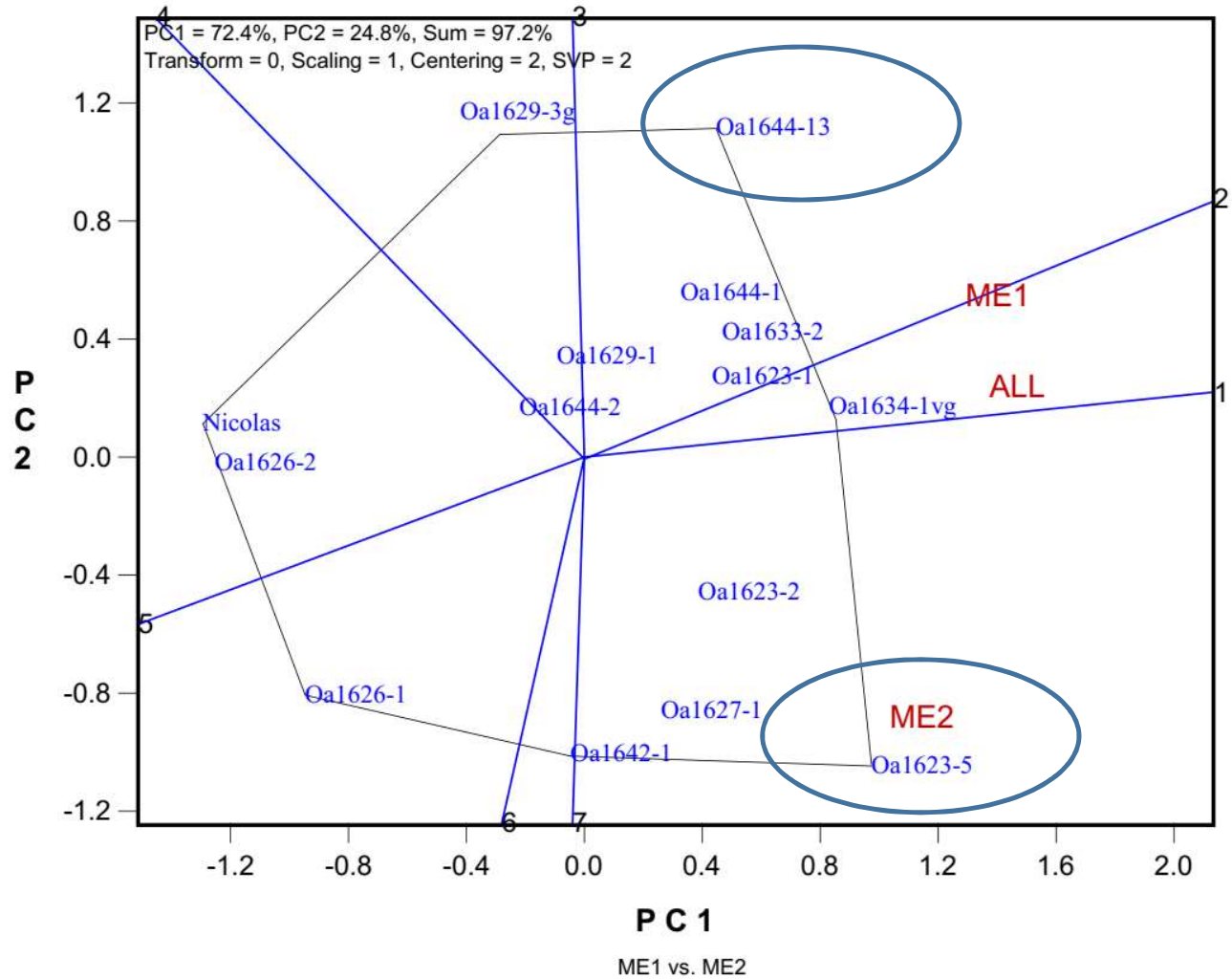
In 2020…

- OA1627-1 was entered into the WCORT (for ME3)
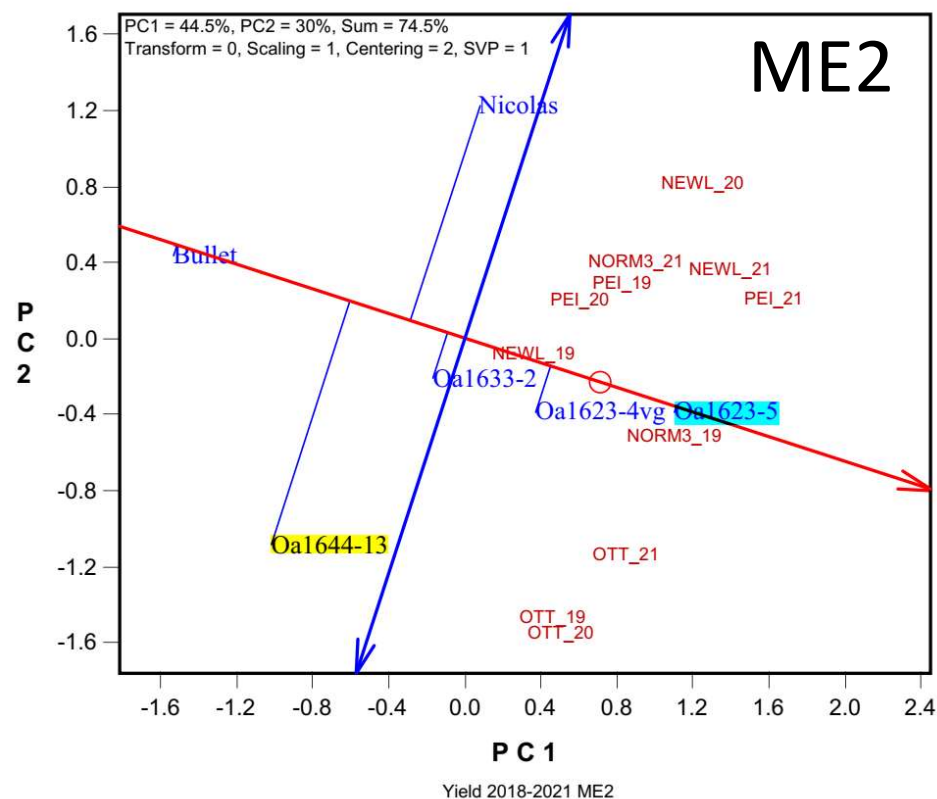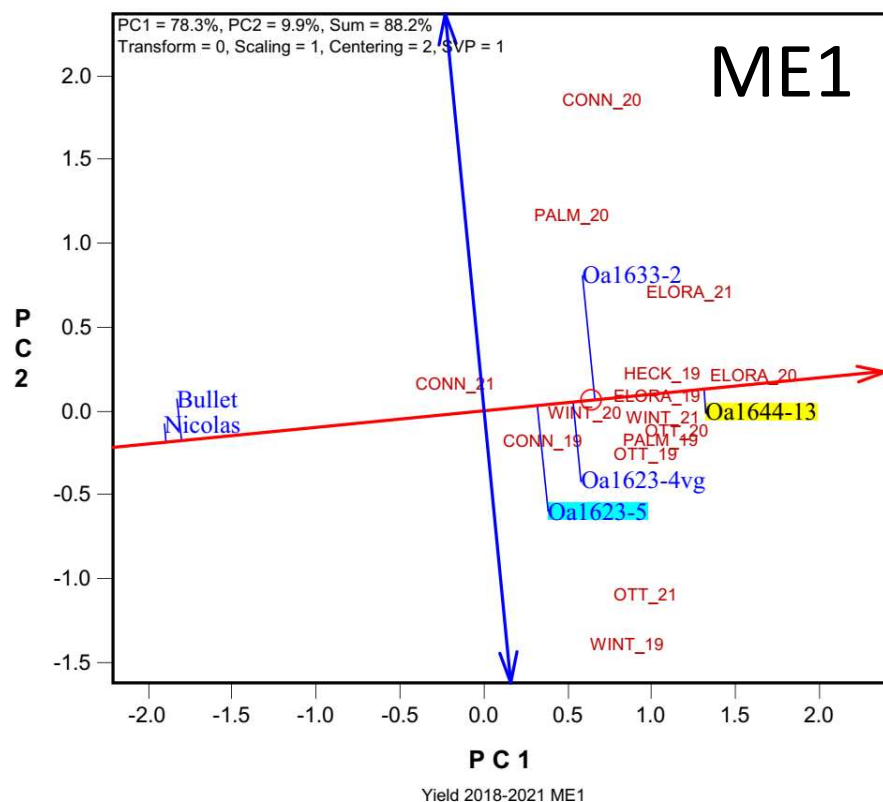- OA1623-5 and OA1644-13 were further tested in ORDC (for ME1 and ME2)
- Two years later…

OA1627-1 was supported for registration for ME3

2020-2021
WCORT data

Entries	Yield (kg/ha)	% of check mean
OA1627-1	5668	106
AC Morgan (CK)	5570	104
Summit (CK)	4963	93
CS Camden (CK)	5545	103

OA1644-13 and OA1623-5 were supported for registration for ME1 and ME2, respectively

PC1 = 78.3%, PC2 = 9.9%, Sum = 88.2%
Transform = 0, Scaling = 1, Centering = 2, SVP = 1

ME1

Yield 2018-2021 ME1

PC1 = 44.5%, PC2 = 30%, Sum = 74.5%
Transform = 0, Scaling = 1, Centering = 2, SVP = 1

ME2

Yield 2018-2021 ME2

It appeared …

- 2 years of nationwide test (c. 10 locations each) was sufficient to identify the highest yielding cultivars for each ME
  - The subsequent registration tests were necessary only to confirm the result and to meet the requirements of the crop recommendation committees

Summary & discussion

- Mega-environment analysis is to reveal <u>repeatable GGE patterns</u> and thereby to divide a target region into meaningful mega-environments

- LG biplot is an effective tool for mega-environment analysis
  - GGE+GGL biplot (Yan 2015) is an alternative

- Three steps to handle GE and improve heritability
  1. Divide the target region into mega-environments
  2. Select and utilize ME-specific cultivars to utilize repeatable GE
  3. Test adequately (locations and years) to accommodate unrepeatable GE

- Three steps to breed ME-specific cultivars: a real story
  1. Develop a hopeful breeding population (east/west crosses)
  2. Preliminary Screening across MEs to identify specific adaptation
  3. Registration Test in the respective ME(s)

- Future approaches
  - Develop ME-specific breeding populations and test within the ME only?
  - Use ME-specific GS models to replace the Preliminary nationwide screening?

Acknowledgement

- The ORDC oat breeding team
  - Brad deHaan, Matt Hayes

- Collaborators conducting the multi-location trials
  - Nathan Mountain, Denis Pageau/Isabelle Morasse/Genevieve Telmosse, Allan Cummiskey/Dan MacEachern, …

- The oat quality team
  - Judith Fregeau-Reid/Mehri Hadinezhad

- The oat disease team
  - Jim Menzies, Allen Xue

- Other Canadian public oat breeding programs
  - Aaron Beattie, Jennifer Mitch-Fetch/Kirby Nilsen, (Mike McMullen)

- The SOO host
  - Charlene Wight, James Bradeen, Pablo Olivera Firpo, Annie Harvieux

OA1644-13 @Ottawa, 2021