

Effect of modelling unknown parent groups on genetic trends of fertility traits

Andres Legarra, andres.legarra@uscdcb.com

Paul VanRaden, paul.vanraden@usda.gov

Interbull, Aug 26 2023



Intro Fertility traits & trends

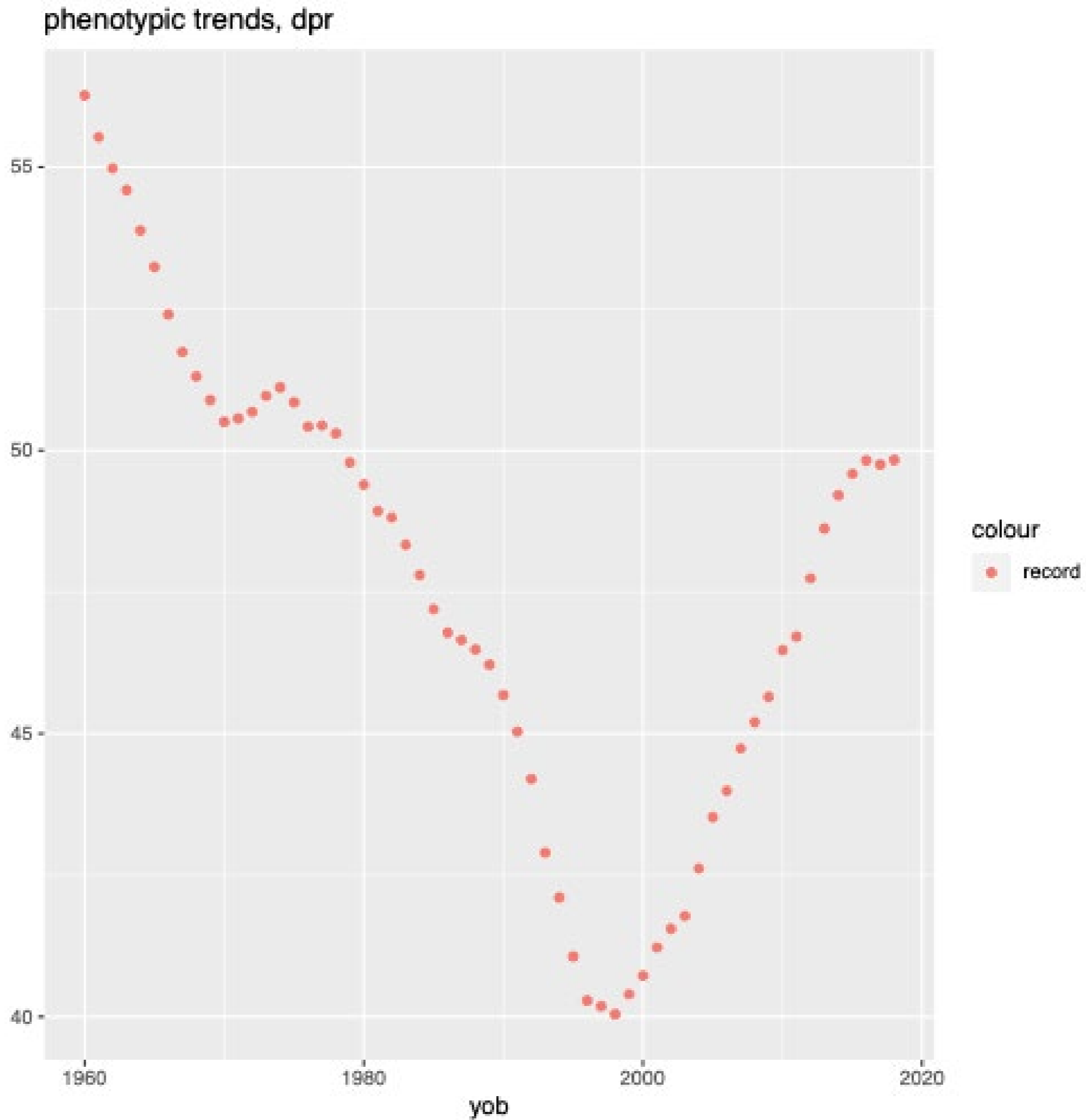
- This is all about traditional (pedigree-based) evaluations: BLUP
- Fertility traits models are complicated because:
 - low heritability (1%) with different data pattern for each trait
 - negative correlation with milk yield => correlated genetic trend => hard to account for
 - correlations and even traits might change with time (natural mating vs. AI; hormonal treatments; heat detection)
 - latest genetic trends also unstable because heifer fertility arrives before cow fertility

Traits

- ccr: Cow Conception Rate (recorded since 2000)
- hcr: Heifer Conception Rate (recorded since 2000)
- dpr: Daughter Pregnancy Rate (recorded since 1960)
- efc: Early First Calving (recorded since 1960)

 Focus is on ccr, dpr with high genetic correlation 0.86

Phenotypic trend (Holstein)



Genetic trends

- Expected ΔG due to correlation with MY, from 1960 to 2000
 - DPR: $\Delta G = -6.86$
 - in 1960-2000 ($\Delta G = 4.2$ SD for MY) * (-0.34 genetic correlation) *(SD = 4.9)
=> $\Delta G = -6.86$
 - CCR: $\Delta G = -5.17$
 - In CCR the period 1960-2000 has no data so the trend is inferred from correlated DPR

Modelling

- We expect to see an initial decrease in fertility due to unfavorable correlation with MY, then an increase as fertility is selected for since ~2000
- However MY is not included in the genetic evaluation for these traits
- Unobserved selection should be accounted for by Unknown Parent Groups (UPG)
- UPGs and herd-year (management trends) have large collinearity
- “Classical” theory of fixed UPG is little satisfying because they don’t correlate across years and they don’t correlate across traits

Models

- Data
 - 4-trait multiple trait BLUP, 94M records, 94M animals in pedigree
 - missing records go from 4% (dpr) to 87% (hcr)
 - all computations on blup90iod3 in ~8h per model, 400 rounds PCG, 800 M equations
 - all-breed evaluation but all results shown are in Holstein
- Models
 - Random UPGs
 - Metafounders based on coancestry (ΔF)
 - Metafounders based on expected genetic change (ΔG)

Random UPG

- Random unknown parent groups, uncorrelated to each other :
 $\text{Var}(\mathbf{g}) = \mathbf{I} \otimes \mathbf{G}_o$ with \mathbf{G}_o a matrix of genetic covariances across traits
- Correlated across traits, but
- uncorrelated across years: each UPG does not “help” its neighbor
- self-relationship of 1 questionable

Metafounders

- $\Gamma \otimes G_o$ measures (ideally) covariances across means of a trait across metafounders
- Covariances across means are described by average relationships (Sorensen & Kennedy 1987):

$$\Gamma = \begin{pmatrix} \bar{A}_0 & \bar{A}_0 & \bar{A}_0 & \dots \\ \bar{A}_0 & \bar{A}_1 & \bar{A}_1 & \dots \\ \bar{A}_0 & \bar{A}_1 & \bar{A}_2 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots \\ 0 & 2t_1\Delta F & 2t_1\Delta F & \dots \\ 0 & 2t_1\Delta F & 2t_2\Delta F & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

- We use this matrix for MF
- This matrix can be set up using inbreeding coefficients ΔF and times
- The inverse is a tri-diagonal sparse matrix where each MF depends on the previous one.
- Matrix Γ was defined within breed in 2 manners

- Metafounder2: Γ within pathway i.e. $\Gamma_{all} = \begin{pmatrix} \Gamma & 0 & 0 & \dots \\ 0 & \Gamma & 0 & \dots \\ 0 & 0 & \Gamma & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$

- Metafounder 1: same Γ across pathways + small number to make pathways different i.e. $\Gamma_{all} = \begin{pmatrix} \Gamma + I\alpha & \Gamma & \Gamma & \dots \\ \Gamma & \Gamma + I\alpha & \Gamma & \dots \\ \Gamma & \Gamma & \Gamma + I\alpha & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$

Metafounders with expected genetic change

- We don't have all selection history yet means have been pulled by selection
- We can sketch a theory about the magnitude of expected genetic change. Due to selection, the mean of the trait

may have shifted from one MF to the next : $\begin{pmatrix} \mu_0 = 0 \\ \mu_1 = 0 + \Delta G_1 \\ \mu_2 = \mu_1 + \Delta G_2 \\ \dots \end{pmatrix}$

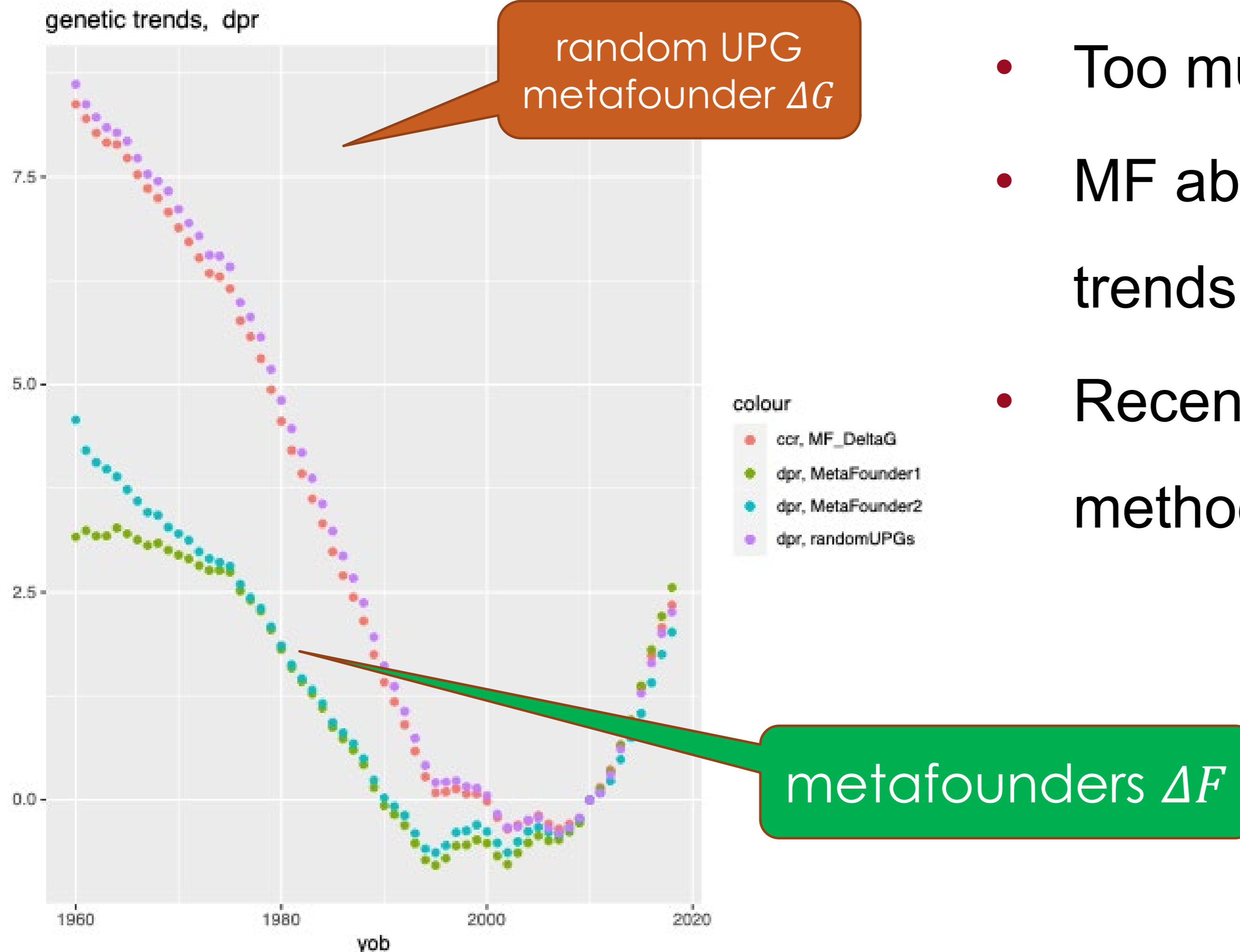
- so $\text{Var}(\mu_1) = \Delta G^2$ (assuming that we don't know the direction of change, $E(\Delta G_1) = 0$). As a result we get something

like this: $\Gamma = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & (\Delta G)^2 & (\Delta G)^2 & (\Delta G)^2 \\ 0 & (\Delta G)^2 & (t_1 \Delta G)^2 & (t_1 \Delta G)^2 \\ 0 & (\Delta G)^2 & (t_1 \Delta G)^2 & (t_2 \Delta G)^2 \end{pmatrix}$

- We put $\Delta G = 0.034 = 1.4/40$ correlated response due to MY per year
- Feels like "cheating" Γ to be more flexible
- This is very crude theory and not sure it's a good idea

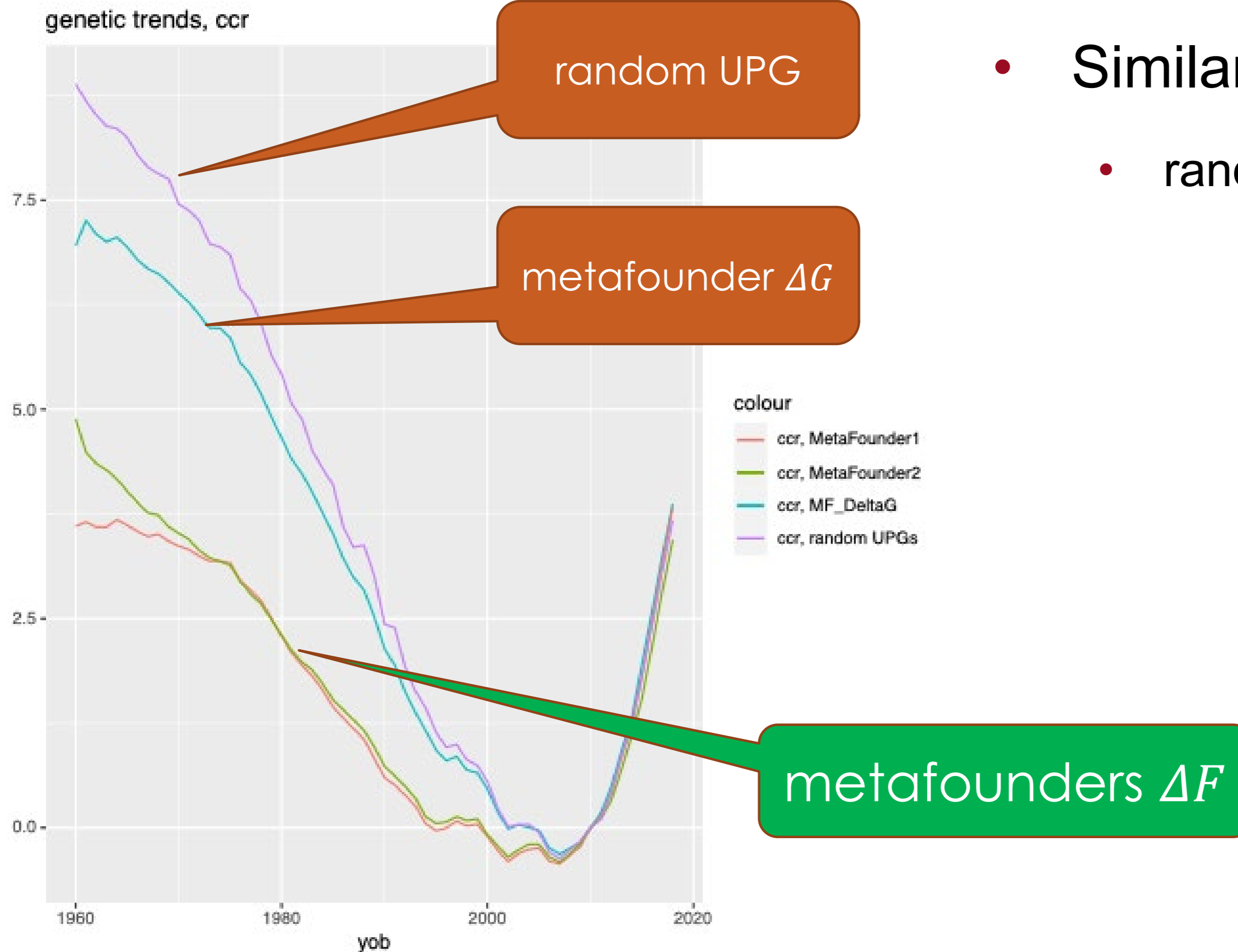
Results

Genetic trends DPR



- Too much change for Random UPG
- MF about right and converge to similar trends
- Recent trends are identical for all methods 😊

Genetic trends CCR



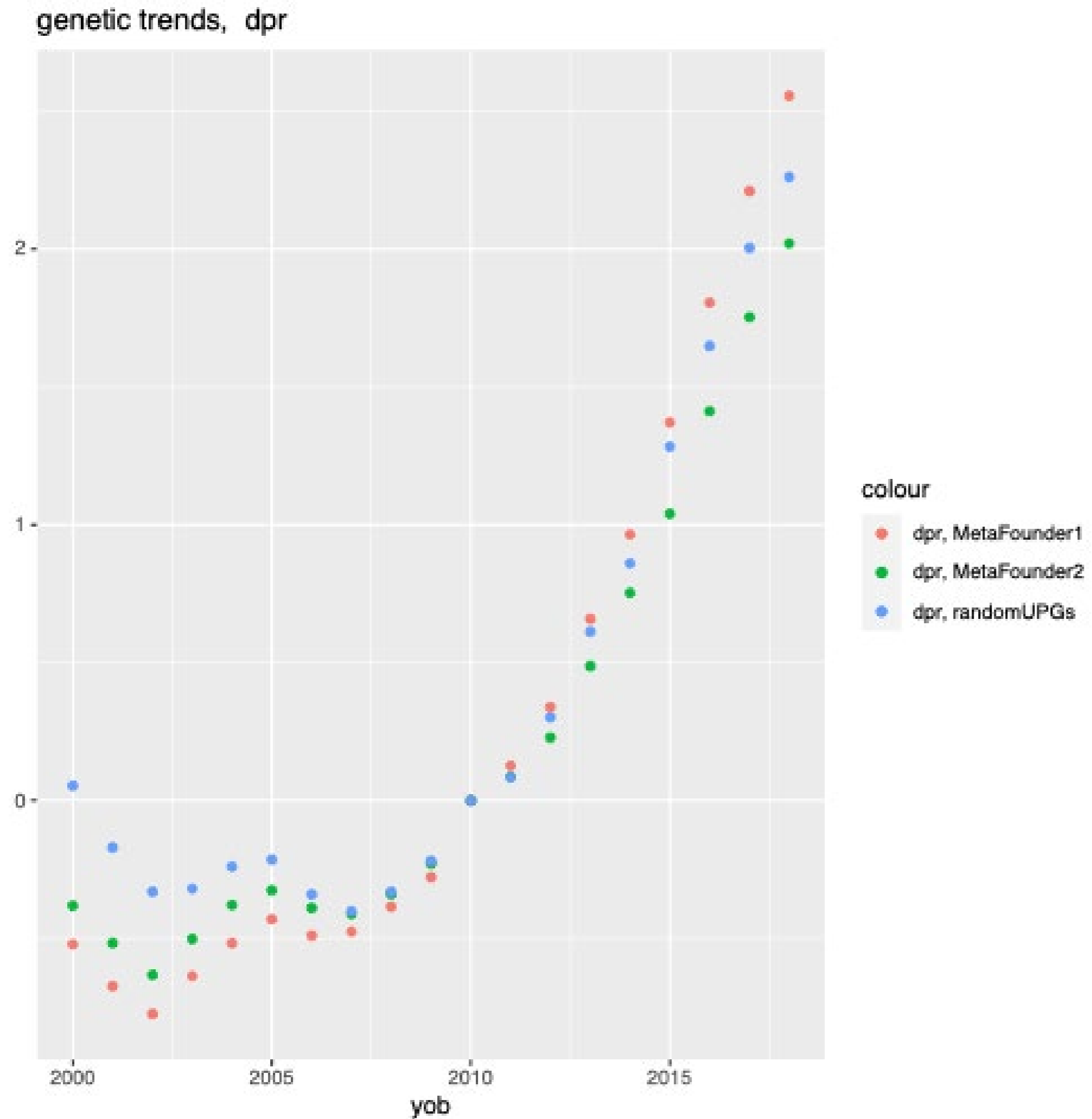
- Similar to DPR as expected
 - random UPG, MF ΔG too much change

Genetic trends

- Change from ~1960 to ~2000
- for ccr, all trend until 2000 is inferred through genetic correlation
- MF is a bit shorter than expected
- Random UPGs & Metafounder ΔG larger than expected

	dpr	ccr
Phenotypic	-16	
Expected genetic	-6.86	-5.17
Random UPG	-9	-9
MF1	-4	-5.3
MF2	-5.2	-5.3
MF_DeltaG	-8.7	-7.6

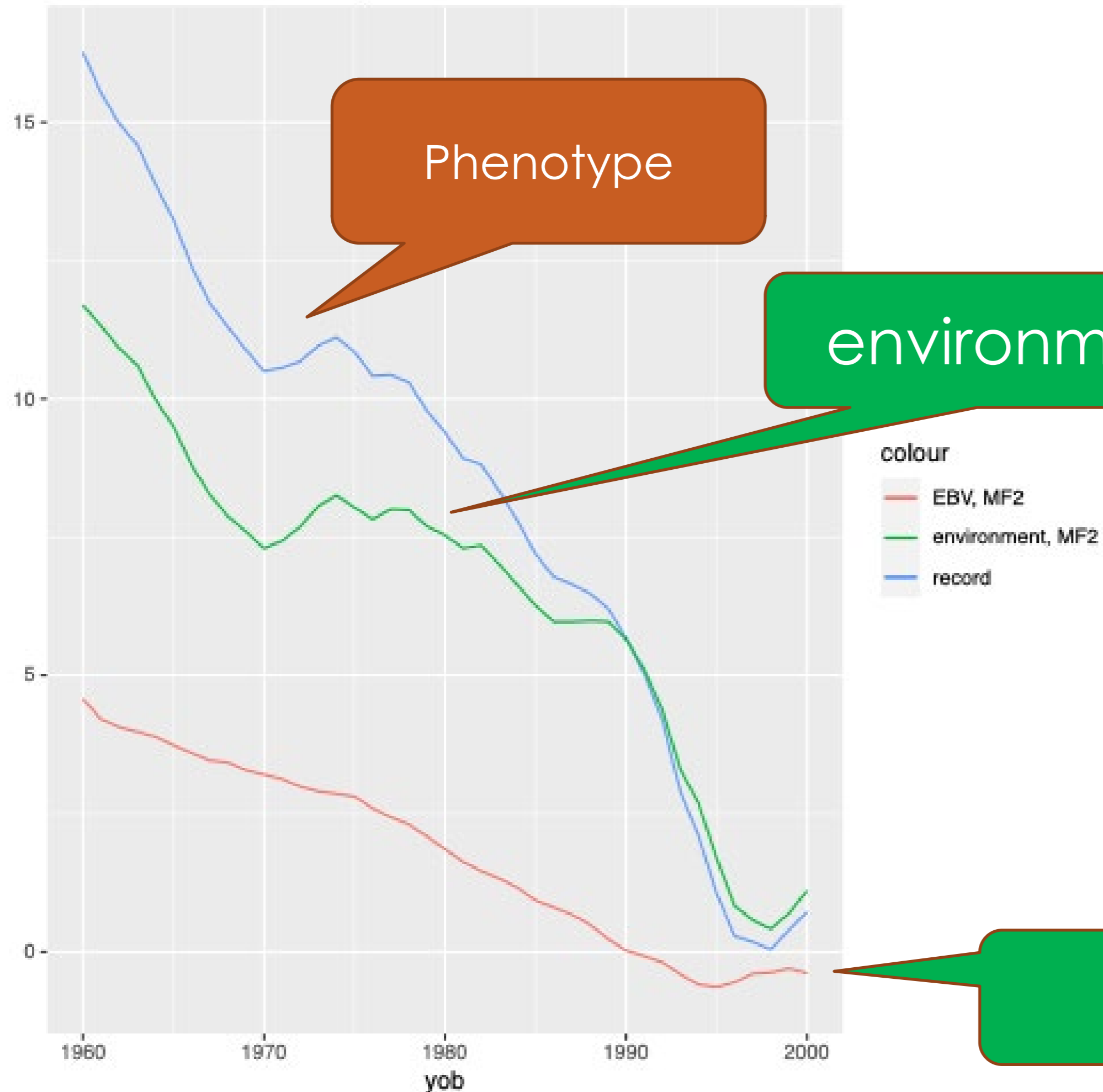
Recent trends DPR



- “fast” change
- some spread across methods

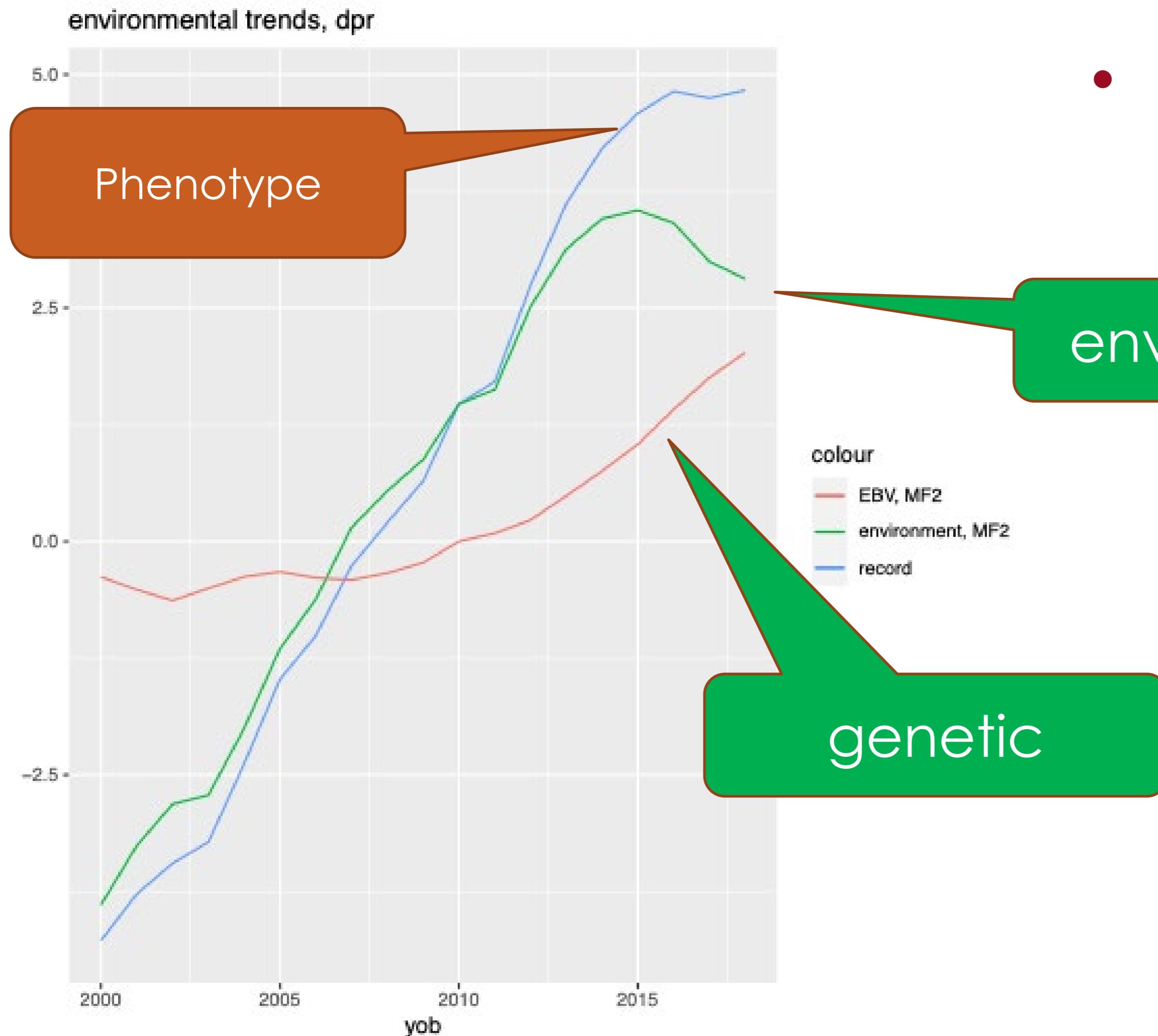
Past environmental vs. genetic trend

environmental trends, dpr



- The negative trend is roughly split into $\frac{2}{3}$ environment and $\frac{1}{3}$ genetics

Recent environmental vs. genetic trend



- The positive trend is roughly split into $\frac{1}{2}$ environment and $\frac{1}{2}$ genetics

Some preliminary conclusions

- How to fit UPGs / metafounders does affect historical genetic trends
- Using some form of prior information for UPGs/MF helps obtaining meaningful genetic trends but which exact prior is hard to say
 - Random UPGs do not work too bad but it's still too much change and the prior is arbitrary
 - The “change in coancestry” (I based on ΔF) is a bit too restrictive but it has a genetic interpretation and is “reproducible” from trait to trait
 - The “expected genetic change” (I based on ΔG) seems too large and it's hard to generalize to other traits => I don't recommend
- We'll see how this interacts with SSGBLUP

Thank you and thanks to all dairy producers

