Correa, K., Bink, M., Vereijken, A., ten Napel, J., Martin, K., Vandenplas, J., Calus, M., Janssen, K. (2026) Accounting for different ploidy levels in trout genetic evaluations in a commercial breeding program, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23702: https://doi.org/10.31274/wcgalp.23702

Accounting for different ploidy levels in trout genetic evaluations in a commercial breeding program

K. Correa[1*], M. Bink[1], A. Vereijken[1], J. ten Napel[2], K. Martin[1], J. Vandenplas[2], M.P.L. Calus[2], and K. Janssen[1]

*[1]Hendrix Genetics, Boxmeer, The Netherlands; [2]Wageningen University & Research, Wageningen, The Netherlands; *Corresponding author: katharina.correa@hendrix-genetics.com*

Introduction

Triploid salmonids are produced at commercial scale to induce sterility, thereby preventing maturation issues during production and reducing ecological risks posed by escapees. Triploidy is routinely induced by hydrostatic pressure shocking eggs, preventing the extrusion of the second polar body shortly after fertilization. Unreduced gametes may originate through first- or second-division restitution (FDR or SDR), corresponding to retention of the first or second polar body, respectively. Triploid induction protocols in salmonids specifically target SDR, leading to retention of sister chromatids during meiosis II (Iqbal *et al*., 2020).

Routine genetic evaluations often rely on the diploid numerator relationship matrix (**A**) built from pedigree data. For autopolyploids or mixed-ploidy pedigrees, diploid computation rules are not appropriate. In these systems, meiosis can generate gametes that carry two or more alleles identical by descent (IBD), and founders may show inbreeding even without consanguinity due to double reduction. Consequently, the self-relatedness terms (A_{ii}) must be parameterized with ploidy-aware rules and must account for the probability λ that two genes drawn at random from a gamete may be IBD as a result of the two copies being inherited from a single parental chromosome (Kerr *et al.*, 2012; Hamilton and Kerr, 2018).

Such structural misspecifications of **A** can in turn affect variance components and induce Estimated Breeding Values (EBV) dispersion/level bias (Hickey *et al*., 2008). A general solution is to first compute diagonal elements and off-diagonal elements between parents of the coancestry matrix (**K**), and secondly compute $\mathbf{A}^{-1}$ through more appropriate scaling, explicitly modelling gametic ploidy and using the appropriate value for λ. This strategy was formalized for autopolyploids by Kerr et al. (2012) and extended to handle multiple ploidy levels (including triploids) and sex-specific IBD by Hamilton and Kerr (2018).

In this study we quantify how ploidy-aware $\mathbf{A}^{-1}$ with different λ-values modifies variance components and the scale and dispersion of EBVs relative to a ploidy-ignored baseline. Finally, we outline operational practices to maintain appropriate EBV scaling in a mixed-ploidy commercial breeding program.

Materials & Methods

Harvest weight in rainbow trout (*Oncorhynchus mykiss*) was analyzed across two year-classes in two different environments. Triploids (n=3,691; all-female), produced by post-fertilization pressure shock, were reared in a river site located in Washington State, United States and assigned to families through SNP markers. Diploids (n=2,405; all female) were PIT-tagged and reared in raceways located in Troutlodge's Eastern Washington farm, U.S. Each year-class consisted of ~150 families. Most families were represented in both the diploid and triploid groups for that year-class, although a small number were not recovered in the triploid harvest.

We fitted a bivariate animal model with traits "triploid harvest weight" and "diploid harvest weight" as two different traits. Fixed effects were pond and year x spawn-week; random effects

were animal, common family environment, and residual. We compared four scenarios, each with a specific A^{-1} matrix, calculated considering ploidy level and λ: (1) baseline (ploidy ignored, implicitly λ=0); (2a) ploidy-aware with λ=0 (no intragametic IBD, the two maternal copies behave as distinct); (2b) λ=0.918 as SDR-plausible reference (unreduced maternal gamete with limited recombination; higher intragametic IBD, as tested in Hamilton and Kerr (2018)); and (2c) λ=1 (maternal double haploid, without crossover). We consider scenario (2b) the most realistic for salmonids because unreduced maternal gametes used for triploid induction arise predominantly via second-division restitution. Under SDR, recombination generates high, but not complete, intragametic identity-by-descent, so λ is expected to be close to 1. This expectation is consistent with the salmonid recombination landscape, characterized by pronounced heterochiasmy and persistent homeologies, which favor limited centromere-proximal recombination in females (Lien *et al*., 2011). Accordingly, λ=0.918 (scenario (2b)) provides a biologically plausible reference value for routine genetic evaluation. Using MiXBLUP (Vandenplas *et al*., 2022), we constructed alternative inverse numerator matrices (A^{-1}) that account for gametic ploidy (τ) and λ, for each of the four scenarios. First, diagonal elements of the coancestry matrix **K** follow the polyploid identity $k_{ii} = [1 + (2v_i - 1)F_i]/(2v_i)$, where v_i is half the ploidy level ($v_i = 1$ for diploids and $v_i = 1.5$ for triploids), and F_i depends on parent-specific gametic contributions (τ_{pi}, τ_{qi}) and λ (Hamilton and Kerr, 2018). Off-diagonal elements were built recursively from parental contribution vectors. Second, A^{-1} was recursively computed using elements of **K** and the scaling factor $\left(2\sqrt{v_x v_y}\right)^{-1}$. These rules accommodate odd ploidy levels, sex-specific- λ, and unequal gametic contributions.

After constructing the A^{-1} matrices for each of the four scenarios, the AI-REML algorithm implemented in DMU v6 (Madsen and Jensen, 2023) was used to estimate variance components and EBVs with their standard errors for both diploid and triploid traits. We supplied scenario-specific A^{-1} as a sparse matrix to DMU before runs. EBVs were centered to the 2016 year-class and scaled by the additive genetic standard deviation (σ_a) to enable cross-scenario comparability. Model fit was evaluated by regressing reference breeding values from scenario (2b) on the corresponding EBVs for the triploid trait to quantify dispersion (slope) and level bias (intercept).

Results

Ignoring ploidy inflated triploid harvest weight h^2 to 0.30 (±0.05), whereas ploidy-aware analyses stabilized triploid harvest weight h^2 to 0.22 (±0.04) across scenarios, regardless of the λ value used (Table 1). Diploid harvest weight h^2 remained 0.48 (±0.04) irrespective of the scenario. The genetic correlation between ploidies stayed ≈0.67 ± 0.1, indicating largely shared additive biology across ploidies once scale is handled correctly. These outcomes are consistent with the expectation that correcting A_{ii} via ploidy and λ primarily results in changes in h^2 while preserving coancestry patterns (Kerr et al., 2012; Hamilton and Kerr, 2018).

Table 1. Genetic parameters and genetic correlations for harvest weight for diploid and triploid weights under different relationship matrix scenarios.

Scenario	Trait	σ^2_a[5]	σ^2_c[6]	σ^2_e[7]	$h^2 \pm SE$[8]	rG[9]
Baseline[1]	Triploid	78,729	11,234	173,889	0.30 ± 0.05	0.65 ± 0.11
	Diploid	31,179	2,098	32,019	0.48 ± 0.04	
2a[2]	Triploid	57,349	5,559	201,740	0.22 ± 0.04	0.68 ± 0.11
	Diploid	31,417	2,028	31,953	0.48 ± 0.04	
2b[3]	Triploid	50,218	7,518	173,897	0.22 ± 0.04	0.67 ± 0.11
	Diploid	31,216	2,049	32,043	0.48 ± 0.04	
2c[4]	Triploid	49,671	7,675	171,741	0.22 ± 0.04	0.67 ± 0.11
	Diploid	31,200	2,047	32,050	0.48 ± 0.04	

[1] Baseline: Conventional diploid relationship matrix ignoring ploidy differences.
[2] Scenario 2a: Ploidy-aware relationship matrix with $\lambda = 0$ (no intragametic IBD).
[3] Scenario 2b: Ploidy-aware relationship matrix with $\lambda = 0.918$ (SDR-plausible reference).
[4] Scenario 2c: Ploidy-aware relationship matrix with $\lambda = 1$ (maternal double haploid).
[5] σ^2_a: Additive genetic variance
[6] σ^2_c: Common environmental variance
[7] σ^2_e: Residual variance
[8] $h^2 \pm SE$: heritability $(\sigma^2_a/(\sigma^2_a + \sigma^2_c + \sigma^2_e)) \pm$ standard error
[9] rG: Genetic correlation between diploid and triploid harvest weight

Pairwise comparisons revealed highly concordant A^{-1} off-diagonals between ploidy-aware scenarios (e.g., r≈0.99 between λ=1 and λ=0.918), but more pronounced differences in the diagonals when contrasting baselines vs. ploidy-aware A^{-1} (r=0.91 for both λ=1 and λ=0.918 vs. baseline; r=0.77 for λ=0 vs. baseline). This pattern confirms that inbreeding scaling—not coancestry—is most sensitive to λ and ploidy encoding, as predicted by the formulation (Kerr et al., 2012; Hamilton and Kerr, 2018).

Using scenario (2b) as reference, and before centering and scaling the EBVs, the baseline diploid model (ignoring ploidy) showed a clear dispersion bias (slope = 0.79; bias = 21.21%). After centering and scaling, the slope improved to 0.97 (bias = 1.35%). The ploidy-aware scenarios yielded slopes near unity after the same step (λ = 0: 1.01; λ = 1: >0.99) with negligible level bias. These findings are consistent with (i) the established bias/dispersion diagnostics—model mis-specification can induce level and dispersion bias in EBVs (Hickey *et al*., 2008); and (ii) the LR-method behaviour reported by Macedo *et al*. (2020), whereby slopes approach 1 under correctly specified models and exhibit bias under mis-specification. In our setting, diploid rules for A are not appropriate in autopolyploids/mixed-ploidy as shown by Kerr et al. (2012) and Hamilton and Kerr (2018), providing a mechanistic rationale for the baseline model's dispersion bias and for the improved calibration under ploidy-aware A^{-1}.

Discussion

The observed sensitivity of triploid h^2 to ploidy and λ arises from how gametic ploidy and λ govern A_{ii} through double reduction and SDR meiosis. Off-diagonals, governed by pedigree paths, remained highly concordant across scenarios; thus, differences concentrate on inbreeding scaling. This is precisely what the derivation predicts for autopolyploids and its extensions to multiple ploidy levels (Kerr *et al*., 2012; Hamilton and Kerr, 2018).

Salmonids show extensive homeologies and marked sex-specific recombination landscapes, including more distal male recombination and broader female maps. Such features support high within-gamete IBD under SDR, making λ values close to one and 0.918 under moderate recombination biologically likely (Lien *et al*., 2011). Our reference scenario (λ=0.918) is

therefore consistent with the species biology rather than ad hoc tuning within the model (Hamilton and Kerr, 2018).

The ploidy-ignored baseline produced over-dispersed EBVs (21.21% bias in slope terms), a hallmark of mis-specified variance structure or relationship scaling. Centre-and-scale corrected most of the dispersion bias (to 1.35%) and yielded negligible level bias, behaviour aligned with the theoretical treatment of bias in EBVs and with LR-based diagnostics that recover slopes close to 1 when models are correct.

For mixed-ploidy breeding programs, we recommend: (i) constructing ploidy-aware A^{-1} that encodes per-parent gametic ploidy and λ, although results suggest relative insensitivity to the value of λ used, and (ii) monitoring the level and dispersion bias at each run. These practices stabilise EBV scales across ploidies and sites and help maintain consistent selection ranks when production relies on triploid grow-outs but selection candidates are diploid broodstock.

Conclusions

Accounting for ploidy and within-gamete IBD in the relationship matrix corrects triploid h^2 inflation and delivers EBVs that are robust across ploidy levels. Thus ploidy-aware A^{-1} provides immediate and practical benefits for commercial trout evaluations across ploidy levels.

Acknowledgements

We thank Troutlodge and colleagues at Hendrix Genetics and Wageningen University & Research for contributions to data collection, genotyping, and analyses.

References

Hamilton M.G., and Kerr R.J. (2018) Theor. Appl. Genet. 131(4):851–860 https://doi:10.1007/s00122-017-3041-y

Hickey J.M., Keane M.G., Kenny D.A., *et al*. (2008) J. Anim. Sci. 86:1047–1056. https://doi:10.2527/jas.2007-0653

Iqbal A., Arab N., Ahmad I., *et al*. (2020). J. Entomol. Zool. Stud. 8(2):381–385.

Kerr R.J., Li L., Tier B., *et al*. (2012) Theor. Appl. Genet. 124:1271–1282. https://doi:10.1007/s00122-012-1785-y

Lien S., Gidskehaug L., Moen T., *et al*. (2011) BMC Genom. 12:615. doi:10.1186/1471-2164-12-615

Macedo F.L., Reverter A., and Legarra A. (2020) J. Dairy Sci 103:529–544. https://doi:10.3168/jds.2019-16603

Madsen P., and Jensen J. (2023) A user's guide to DMU: A package for analysing multivariate mixed models (Version 6, release 5.5). Center for Quantitative Genetics and Genomics, Aarhus University, Tjele, Denmark.

Vandenplas J., Veerkamp R.F., Calus M.P.L., *et al*. (2022) Proc. 12th World Congress on Genetics Applied to Livestock Production (WCGALP), Rotterdam, The Netherlands.