**Table S4 Selection accuracy by scenario, relative cost of phenotyping to genotyping ($P:$G), and the availability of an initial training population**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | With initial training population | | | Without initial training population | | |
| Scenario | $P:$G=1:2 | $P:$G=1:1 | $P:$G=2:1 | $P:$G=1:2 | $P:$G=1:1 | $P:$G=2:1 |
| Male candidates | | | | | | |
| C11, S1 | 0.370.04a ,A | 0.370.04a,A | 0.370.04a,A | 0.370.04a,A | 0.370.04a,A | 0.370.04a,A |
| C11, S2 | 0.940.01b,A | 0.940.01b,A | 0.940.01b,A | 0.940.01b,A | 0.940.01b,A | 0.940.01b,A |
| G10 | 0.890.03c,A | 0.900.02bc,AB | 0.910.01bc,B | 0.810.03b,A \* | 0.840.01b,B \* | 0.870.01b,C \* |
| G9 | 0.900.03bc,A | 0.910.02bc,A | 0.910.01bc,A | 0.850.02c,A \* | 0.870.01bc,B \* | 0.900.01bc,C \* |
| G8 | 0.910.01bc,A | 0.910.01bc,A | 0.910.01bc,A | 0.860.01cd,A \* | 0.890.01c,B \* | 0.900.01bc,B |
| G5 | 0.910.01bc,A | 0.910.00bc,A | 0.910.01bc,A | 0.900.01d,A | 0.910.01c,A | 0.910.01c,A |
| G2 | 0.910.01bc,A | 0.910.00bc,A | 0.900.01bc,A | 0.900.01d,A | 0.900.01c,A | 0.900.01bc,A |
| G1 | 0.890.01c,A | 0.900.01c,A | 0.890.01c,A | 0.890.01cd,A | 0.890.01c,A | 0.890.01bc,A |
| Sires | | | | | | |
| C11 | 0.860.05a,A | 0.860.05a,A | 0.860.05a,A | 0.860.05a,A | 0.860.05a,A | 0.860.05a,A |
| G10 | 0.750.04b,A | 0.750.03b,A | 0.730.05b,A | 0.670.08bc,A \* | 0.680.05cde,A \* | 0.670.06b,A \* |
| G9 | 0.760.04b,A | 0.720.06bc,AB | 0.690.05c,A | 0.700.05b,A \* | 0.720.05bc,A | 0.710.05b,A |
| G8 | 0.760.03b,A | 0.690.05cd,B | 0.680.06c,B | 0.710.05b,A \* | 0.740.05b,A \* | 0.700.07b,A |
| G5 | 0.680.07c,A | 0.670.08de,A | 0.690.04c,A | 0.680.05bc,A | 0.690.05cd,A | 0.690.03b,A |
| G2 | 0.670.05c,A | 0.670.05de,A | 0.670.04c,A | 0.650.06c,A | 0.640.07e,A | 0.690.05b,A |
| G1 | 0.660.06c,A | 0.630.05e,A | 0.670.04c,A | 0.670.04bc,A | 0.670.03de,A | 0.690.05b,A |
| Female candidates | | | | | | |
| C11 | 0.450.02a,A | 0.450.02a,A | 0.450.02a,A | 0.450.02a,A | 0.450.02a,A | 0.450.02a,A |
| G10 | 0.480.01ab,A | 0.480.01ab,A | 0.510.01b,B | 0.460.02ab,A \* | 0.470.02ab,AB | 0.490.01b,B \* |
| G9 | 0.490.02b,A | 0.500.01b,B | 0.520.01b,C | 0.470.02ab,A \* | 0.490.02bc,B | 0.520.01bc,C |
| G8 | 0.510.01b,A | 0.510.01b,A | 0.540.01bc,B | 0.490.02bc,A \* | 0.520.01cd,B | 0.530.01cd,C |
| G5 | 0.510.01bc,A | 0.550.01c,B | 0.570.01c,C | 0.520.01cd,A | 0.550.01de,B | 0.570.01d,C |
| G2 | 0.550.01cd,A | 0.570.01c,B | 0.570.01c,B | 0.550.01d,A | 0.560.02e,AB | 0.570.01d,B |
| G1 | 0.560.01d,A | 0.560.01c,A | 0.560.01c,A | 0.550.01d,A | 0.560.01e,A | 0.560.01d,A |
| Cows | | | | | | |
| C11 | 0.480.03a,A | 0.480.03a,A | 0.480.03a,A | 0.480.03a,A | 0.480.03a,A | 0.480.03a,A |
| G10 | 0.560.02b,A | 0.590.02b,B | 0.630.01b,C | 0.530.01b,A \* | 0.560.01b,B \* | 0.610.01b,C \* |
| G9 | 0.590.03bc,A | 0.630.02c,B | 0.700.01c,C | 0.570.02bc,A \* | 0.620.02c,B | 0.680.02c,C \* |
| G8 | 0.620.02c,A | 0.670.02c,B | 0.740.02d,C | 0.600.02c,A \* | 0.660.01d,B | 0.730.02d,C |
| G5 | 0.700.02d,A | 0.770.01d,B | 0.790.02e,C | 0.690.02d,A | 0.760.01e,B | 0.780.02e,B |
| G2 | 0.760.02e,A | 0.790.02d,B | 0.780.01e,AB | 0.760.01e,A | 0.770.02e,A \* | 0.770.01de,A |
| G1 | 0.770.02e,A | 0.770.02d,A | 0.770.01de,A | 0.760.01e,A | 0.760.02e,A | 0.760.02de,A |

\*The table presents the means and standard deviations (subscript) across 10 replicates for the conventional (C) and genomic (G) scenarios, with numbers indicating the number of phenotype records per lactation. Conventional selection implemented two-stage selection for males, hence we present the accuracy of pre-selection for progeny testing (S1) and the accuracy sire selection (S2). In genomic scenarios the male candidates were genotyped and non‑phenotyped. We also present the accuracy for sires currently used in artificial insemination (sires), for non‑genotyped and non‑phenotyped females (female candidates), and for all active phenotyped cows and bull dams (cows). Lower-case letters denote statistically significant differences between scenarios within the same $P:$G and upper-case letters between different $P:$G within the same scenario. Stars denote statistically significant difference between corresponding scenarios with and without an initial training population.