



HAL
open science

Research note: double digest Restriction site Associated DNA sequencing is a suitable alternative to high density SNP chips for genomic selection of layers

Mathilde Doublet, Lorry Bécot, Frédéric Lecerf, Sophie Allais

► **To cite this version:**

Mathilde Doublet, Lorry Bécot, Frédéric Lecerf, Sophie Allais. Research note: double digest Restriction site Associated DNA sequencing is a suitable alternative to high density SNP chips for genomic selection of layers. Poultry Science, 2025, 104 (7), pp.105115. <10.1016/j.psj.2025.105115>. <hal-05019793>

HAL Id: hal-05019793

<https://hal.science/hal-05019793v1>

Submitted on 22 May 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



Research note: double digest restriction site associated DNA sequencing is a suitable alternative to high density SNP chips for genomic selection of layers

Mathilde Doublet^a, Lorry Bécot^b, Frédéric Lecerf^{a,1}, Sophie Allais^{a,1,*}

^a PEGASE, INRAE, Institut Agro 35590, Saint Gilles, France

^b NOVOGEN, 5 rue des Compagnons 22960 Plédran, France

ARTICLE INFO

Keywords:

ddRAD-seq

Genomic evaluation

Imputation

Laying hens

ABSTRACT

For the genomic selection of laying hens, the costs of genotyping are still a challenge. Alternative genotyping approaches as low density SNP chips have been developed by breeding companies to avoid the use of the high-density SNP chip (HD). Another approach called **double digest restriction site associated DNA sequencing** (ddRAD-seq), associated with an appropriate quality control was also proven to be reliable to genotype laying hens. From this perspective, the use of ddRAD-seq genotypes as an alternative to HD SNP chip genotypes was studied for the genomic evaluation of 179 candidate individuals in a commercial pure line of laying hens of Rhode Island.

Genomic evaluations based on ancestry were performed in three different scenarios and for four egg quality traits. The first scenario is the reference scenario using only HD SNP chip genotypes. The **genomic estimated breeding values** (GEBV) were calculated for the 179 candidates with 437 individuals as the reference population (GEBV_{HDchip437}). In the second scenario, the reference population included only the 50 individuals with true ddRAD-seq genotypes and GEBV_{ddRAD-seq50} were calculated for the 179 candidates, based on their ddRAD-seq genotypes. Finally, for the third scenario of genomic evaluation, the reference population was constituted of 437 individuals, with simulated ddRAD-seq genotypes. The GEBV_{ddRAD-seq437} were calculated for the 179 candidates based on their true ddRAD-seq genotypes. Then, Spearman correlations were calculated between GEBV_{ddRAD-seq50} or GEBV_{ddRAD-seq437} and the reference GEBV_{HDchip437}.

The mean rank correlation between GEBV_{HDchip437} and GEBV_{ddRAD-seq50} was 0.92. The mean correlation between GEBV_{HDchip437} and GEBV_{ddRAD-seq437} was 0.98. The reduced reference population of the GEBV_{ddRAD-seq50} scenario didn't permit for a good representation of the genetic diversity in the population. But, with an equivalent reference population, genomic evaluation based on ddRAD-seq genotypes achieves similar results as genomic evaluation based on HD SNP chip genotypes.

Introduction

Genomic selection, enabled by the development of Next generation sequencing approaches, was first described by [Meuwissen et al., 2001](#). It is a method that aims to estimate the genomic values of the selection candidates, with or without own phenotypes but with own genotypes, based on a reference population having both own genotypes and individual or offspring phenotypes. In the laying hens sector, genomic selection has been enabled by the development of 600 K **high density (HD)** genotyping chip developed by [Kranis et al., 2013](#) and

commercialized by Affymetrix® Axiom®. Genomic evaluation has already been proven more precise than pedigree-based evaluation for laying hens selection ([Picard Druet et al., 2020](#)). One of the main challenges of implementing genomic selection for poultry is to limit the costs while ensuring a good quality of genotyping, as the precision of genomic evaluations depends on the accuracy of genotypes. To implement genomic selection, the cost of HD SNP chip is still a problem for the breeding companies. Then, alternative approaches for laying hens genotyping have been developed. Usually, breeding companies develop **low density (LD)** chips, selecting only a subset of markers from the HD SNP

* Corresponding author at: 65 rue de Saint Briec, Bâtiment 15 35042, Rennes.

E-mail address: sophie.allais@institut-agro.fr (S. Allais).

¹ The authors Lecerf and Allais contributed equally to the work

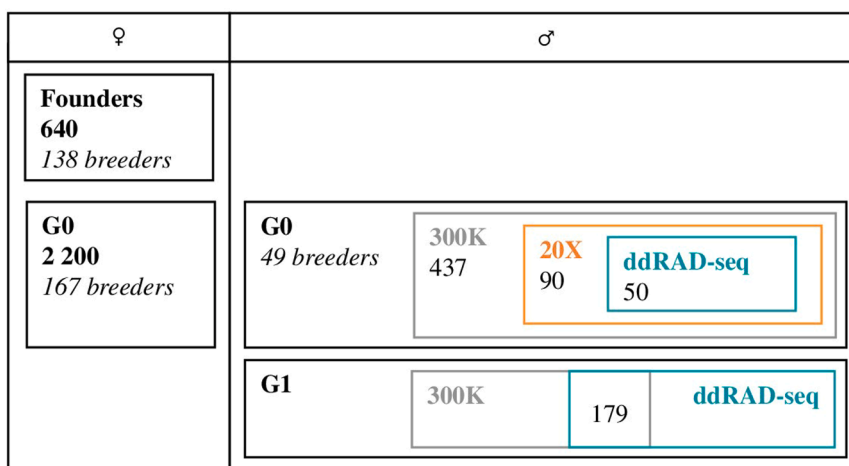


Fig. 1. The schematic diagram of population structure describing phenotyped and genotyped individuals.

chip. In 2024, Doublet et al. showed that **double digest Restriction site Associated DNA sequencing (ddRAD-seq)** could be a reliable genotyping method for laying hens. Compared to LD SNP chips, ddRAD-seq sequencing enables the integration of the whole genome including all nowadays known micro-chromosomes and new markers unavailable on the HD SNP chip. Also, genomic evaluations using simulated ddRAD-seq genotypes were carried by Herry et al., 2023 and showed that ddRAD-seq could be an interesting alternative approach to LD SNP chips for genomic selection.

Then, the aim of this study was to assess the reliability of ddRAD-seq genotypes for the genomic evaluations of laying hens in comparison with HD SNP chip genotypes.

Material and methods

Ethics approval

All blood samples were carried out as part of the commercial and selection activities of Novogen (Plédran, France). These animals studied and the scientific investigations described herein are therefore not to be considered as experimental animals per se, as defined in EU directive 2010/63 and subsequent national application texts. Consequently, we did not seek ethical review and approval of this study as regarding the use of experimental animals. All animals were reared in compliance with national regulations pertaining to livestock production and according to procedures approved by the French Veterinary Services.

Animals and population structure

All animals consisted in a commercial pure line of laying hens of Rhode Island. This line was created and selected by Novogen. The population studied was constituted of 3 456 individuals. The G0 group counted 2 200 females and 437 males. Among the G0 individuals, 167 females and 49 males were breeders for the G1 group. The other individuals of the G0 group were contemporary relatives of the breeders. The G1 group was constituted of 179 male individuals (Fig. 1). A total of 640 females from generations preceding G0 (Founders) were considered. Among them, 138 were breeders for the G0 group. The other female individuals of the founder group were contemporary relatives of the 138 breeders.

Genotyping

All male individuals from G0 and G1 were genotyped using the 600 K Affymetrix® Axiom® HD SNP chip (Kranis et al., 2013). Genotyping acquisition was detailed in Herry et al. (2018). Finally, 301 598 SNP

were retained after quality control. Among the 437 G0 males, 90 were genotyped using 20X **whole genome sequencing (WGS)** as described in Herry et al. (2023), using the 7th version of the reference genome for layers (GRCg7b, Warren et al., 2023). The 20X WGS genotypes were filtered in successive steps with VCFtools, including SNP call rate (> 80 %) and mean depth of sequencing (>5X). A total of 9 256 065 biallelic SNP distributed on chromosomes 1 to 39 and Z were retained. Among the 90 G0 males genotyped with both HD SNP chip and 20X WGS, 50 were genotyped using ddRAD-seq as described in Doublet et al. (2024) using the restriction enzymes Taq1-v2 and Pst1-HF (New England Biolabs, Massachusetts, USA). The ddRAD-seq genotypes were filtered in successive steps with VCFtools, including SNP call rate (> 80 %) and mean depth of sequencing (>5X). A total of 109 752 SNP distributed on chromosomes 1 to 39 and sex chromosome Z was conserved. The 179 G1 male individuals were genotyped using both HD SNP chip and ddRAD-seq, in the same conditions as for the G0 individuals.

Simulated ddRAD-seq

Based on the shared SNP between HD SNP chip and 20X WGS, 347 G0 male individuals genotyped with only HD SNP chip were imputed from 85 other G0 individuals, genotyped with both HD SNP chip and 20X WGS. FImpute V3 (Sargolzaei et al., 2014) was used to impute the 347 G0 individuals. Imputation was performed chromosome by chromosome. The 5 remaining G0 individuals having both 20X WGS and HD SNP chip genotypes were imputed the same way based on their HD genotypes, constituting the validation population for imputation accuracy. For each 5 individuals having both genotypes resulting from the imputation and true 20X WGS genotypes, imputation accuracy was estimated as the mean Pearson correlation, for each chromosome, between true and imputed genotypes. Then, the either true or imputed 9 127 276 SNP of the 437 G0 male individuals were filtered based on the 103 398 common SNP between 20X WGS and ddRAD-seq. For each of the 50 individuals having both simulated (resulting from filtration of 20X WGS) and true ddRAD-seq genotypes (Fig 1), ddRAD-seq simulation accuracy was estimated as the mean Pearson correlation between true and imputed filtered genotypes, per individual. The formatting of the data was performed using Plink and calculations were performed using R. All VCFtools, Plink and R versions were described in Doublet et al. (2024).

Phenotypes

Four traits of egg quality were measured: the **egg weight (EW)**, the **eggshell color (ESC)**, the **eggshell strength (ESS)** and the **albumin height (AH)**. The traits were recorded individually for all 2 840 females

Table 1

Spearman correlations between reference $GEBV_{HDchip437}$ and $GEBV_{ddRAD-seq50}$ and between $GEBV_{HDchip437}$ and $GEBV_{ddRAD-seq437}$. EW: egg weight; ESC: egg shell color; ESS: egg shell strength; AH: Albumin height

Reference scenario	Test scenario	EW	ESC	ESS	AH	Mean corr.	Median corr.	SD corr.
$GEBV_{HDchip437}$	$GEBV_{ddRAD-seq50}$	0.92	0.93	0.84	0.92	0.90	0.92	0.04
$GEBV_{HDchip437}$	$GEBV_{ddRAD-seq437}$	0.98	0.98	0.97	0.97	0.97	0.97	0.005

reared in individual cages, aged between 60 and 90 weeks. Collection of egg and data treatment protocol were described in [Herry et al. \(2023\)](#).

Genomic evaluations

The four traits (EW, ESC, ESS, AH) were evaluated for the 179 G1 candidates using single-step GBLUP methodology implemented in BLUPF90 programs ([Miszta et al., 2002](#)), and with a classical multi-trait animal model. All available female phenotypes were included in the evaluations. The fixed and covariable environmental effects of the model were described in [Picard Druet et al. \(2020\)](#). Genomic evaluations based on ancestry were performed in three different scenarios. The first scenario is the reference scenario using only HD SNP chip genotypes. The **genomic estimated breeding values (GEBV)** were calculated for the 179 G1 candidates with the 437 G0 individuals as the reference population ($GEBV_{HDchip437}$). In the second scenario, all the individuals (reference and candidate population) had true ddRAD-seq genotypes. The reference population included only the 50 G0 individuals with true ddRAD-seq genotypes and $GEBV_{ddRAD-seq50}$ were calculated for the 179 G1 candidates, based on their ddRAD-seq genotypes. Finally, for the third scenario of genomic evaluation, the reference population was constituted of 437 G0 individuals, with 347 having simulated (imputed from HD SNP chip to 20X WGS and filtered) ddRAD-seq genotypes and 90 having simulated (filtered from 20X WGS) ddRAD-seq genotypes. The $GEBV_{ddRAD-seq437}$ were calculated for the 179 G1 candidates based on their true ddRAD-seq genotypes.

The reliability of ddRAD-seq as a genotyping approach for genomic evaluation was assessed by comparing GEBV obtained with true ($GEBV_{ddRAD-seq50}$) or simulated ddRAD-seq ($GEBV_{ddRAD-seq437}$) and reference GEBV obtained with HD SNP chip ($GEBV_{HDchip437}$) using Spearman correlation.

Results and discussion

The aim of this study was to assess the reliability of ddRAD-seq as a genotyping method for ancestry genomic evaluation of laying hens, as an alternative to HD SNP chips. First, only true ddRAD-seq genotypes were used to calculate the $GEBV_{ddRAD-seq50}$ of the 179 G1 candidates, based on a reference population of 50 G0 individuals. The $GEBV_{ddRAD-seq50}$ were compared to their reference $GEBV_{HDchip437}$. Spearman correlations between $GEBV_{ddRAD-seq50}$ and $GEBV_{HDchip437}$ were 0.92 for EW, 0.93 for ESC, 0.84 for ESS and 0.92 for HA ([Table 1](#)) with a mean correlation of 0.90, median correlation of 0.92 and SD for correlations of 0.04.

Even if the correlations were promising, there was still a substantial difference between the rankings of the $GEBV_{ddRAD-seq50}$ and the reference $GEBV_{HDchip437}$. Also, the results indicated that correlations for ESS were lower than for the other traits. It is known in the literature that the size of the reference population directly affects the accuracy of genomic evaluations, as a larger population allows for better capture of genetic diversity and an improved estimation of marker effects on the quantitative traits. Therefore, we hypothesized that the difference between the rankings of $GEBV_{ddRAD-seq50}$ and $GEBV_{HDchip437}$ and the lower correlation for ESS were due to the reduced size of the reference population. So, to compare results of genomic evaluations, it appeared necessary to perform genomic evaluations based on the exact same reference population between the reference HD SNP chip scenario and the ddRAD-seq scenario.

Consequently, the $GEBV_{ddRAD-seq437}$ were calculated using the 437 G0 male individuals as the reference population. But, in the G0 group, only a subset had true ddRAD-seq genotypes. So, ddRAD-seq genotypes were simulated for the 437 G0 male individuals. First, the imputation of the 352 G0 male individuals that didn't have 20X WGS genotypes was performed based on the 298 745 shared SNP between HD SNP chip and 20X WGS. The SNP were distributed on chromosomes 1 to 15, 17 to 28 and Z. On those chromosomes, the 85 G0 individuals of the reference population for imputation had 9 127 276 SNP. The imputation accuracy was assessed for the validation population (5 individuals) having both imputed and true 20X WGS genotypes. The mean correlation between the 9 127 276 imputed and true genotypes was 0.95 with a median correlation of 0.97 and SD of 0.05. The results of imputation were consistent with the results obtained by [Herry et al., 2023](#). Then, to simulate ddRAD-seq genotypes for all the 437 G0 male individuals, the 9 127 276 true or imputed genotypes were reduced to the common SNP with ddRAD-seq. After filtration, 103 398 SNP were retained for each of the 437 individuals. The SNP were distributed on chromosomes 1 to 15, 17 to 28 and Z. For the 50 G0 male individuals having true and simulated ddRAD-seq genotypes, the mean correlation between the 103 398 true and simulated ddRAD-seq genotypes was 0.97. The median correlation was 0.97 and the SD was 0.007. On the common chromosomes (1-15, 17-28 and Z) between true and simulated datasets, simulated ddRAD-seq genotypes were then similar to real ddRAD-seq genotypes. So, simulated ddRAD-seq genotypes for the 437 G0 individuals were included in the calculation of $GEBV_{ddRAD-seq437}$ of the 179 G1 male candidates.

The Spearman correlations between $GEBV_{ddRAD-seq437}$ and $GEBV_{HDchip437}$ were 0.98 for both EW and ESC, and 0.97 for both ESS and HA with a mean correlation of 0.97, median correlation of 0.97 and SD for correlations of 0.005 ([Table 1](#)). It's noteworthy that including simulated ddRAD-seq genotypes in the evaluation for the reference population, implied to exclude genotypes on chromosomes 16 and 29 to 39. Even though it is known that these micro chromosomes carry a significant part of the genetic diversity in the layers population ([Degalez et al., 2024](#)) the results of genomic evaluations were satisfying and constant between traits. This showed that the lower correlation obtained for ESS between the $GEBV_{ddRAD-seq50}$ and $GEBV_{HDchip437}$ was due to the reduced size of the reference population for the calculation of $GEBV_{ddRAD-seq50}$.

In line with the results of [Doublet et al. \(2024\)](#), it appears that ddRAD-seq with an appropriate quality control is a reliable tool for the evaluation of laying hens. To pursue this work, it would be interesting to assess the relative accuracy of ancestry genomic evaluations performed with ddRAD-seq genotypes compared to genomic evaluations including a maximum of genotypes and phenotypes on ancestry, contemporary relatives and offspring. This way, it would be possible to estimate the added value of the missing 16 and 29 to 39 micro chromosomes on the accuracy of genomic evaluation based on ddRAD-seq, compared to HD SNP chip genotypes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research project was supported by the French National Research Agency ANR, within the framework of project ANR-10-GENOM_BTV-015 UtOpIGe, and by the French Institut national de la recherche agronomique et de l'environnement (INRAE), within the framework of the SelGen metaprogram. Mathilde Doublet is a PhD fellow supported by the French Institut national de la recherche agronomique et de l'environnement (INRAE) and L'institut Agro Rennes-Angers.

References

- Degalez, F., Charles, M., Foissac, S., Zhou, H., Guan, D., Fang, L., Klopp, C., Allain, C., Lagoutte, L., Lecerf, F., Acloque, H., Giuffra, E., Pitel, F., Lagarrigue, S., 2024. Enriched atlas of lncRNA and protein-coding genes for the GRCg7b chicken assembly and its functional annotation across 47 tissues. *Sci. Rep.* 14, 6588. <https://doi.org/10.1038/s41598-024-56705-y>.
- Doublet, M., Degalez, F., Lagarrigue, S., Lagoutte, L., Gueret, E., Allais, S., Lecerf, F., 2024. Variant calling and genotyping accuracy of ddRAD-seq: comparison with 20X WGS in layers. *PLOS ONE* 19, e0298565. <https://doi.org/10.1371/journal.pone.0298565>.
- Herry, F., Hérault, F., Lecerf, F., Lagoutte, L., Doublet, M., Picard-Druet, D., Bardou, P., Varenne, A., Burlot, T., Roy, P.L.e, Allais, S., 2023. Restriction site-associated DNA sequencing technologies as an alternative to low-density SNP chips for genomic selection: a simulation study in layer chickens. *BMC Genom.* 24, 271. <https://doi.org/10.1186/s12864-023-09321-5>.
- Herry, F., Hérault, F., Picard Druet, D., Varenne, A., Burlot, T., Le Roy, P., Allais, S., 2018. Design of low density SNP chips for genotype imputation in layer chicken. *BMC Genet.* 19, 108. <https://doi.org/10.1186/s12863-018-0695-7>.
- Kranis, A., Gheyas, A.A., Boschiero, C., Turner, F., Yu, L., Smith, S., Talbot, R., Pirani, A., Brew, F., Kaiser, P., Hocking, P.M., Fife, M., Salmon, N., Fulton, J., Strom, T.M., Haberer, G., Weigend, S., Preisinger, R., Gholami, M., Qanbari, S., Simianer, H., Watson, K.A., Woolliams, J.A., Burt, D.W., 2013. Development of a high density 600K SNP genotyping array for chicken. *BMC Genom.* 14, 59. <https://doi.org/10.1186/1471-2164-14-59>.
- Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 And related programs (BGF90) 2.
- Picard Druet, D., Varenne, A., Herry, F., Hérault, F., Allais, S., Burlot, T., Roy, P.L.e, 2020. Reliability of genomic evaluation for egg quality traits in layers. *BMC Genet.* 21, 17. <https://doi.org/10.1186/s12863-020-0820-2>.
- Sargolzaei, M., Chesnais, J.P., Schenkel, F.S., 2014. A new approach for efficient genotype imputation using information from relatives. *BMC Genom.* 15, 478. <https://doi.org/10.1186/1471-2164-15-478>.
- Warren, W., Fedrigo, O., Tracey, A., Mason, A.S., Formenti, G., Francesco, P., Wu, Z., Murphy, T., Schneider, V., Stiers, K., Rice, E.S., Coghill, L.M., Anthony, N., Okimoto, R., Carroll, R., Mountcastle, J., Balacco, J., Haase, B., Yang, C., Zhang, G., Smith, J., Drechsler, Y., Cheng, H.H., Howe, K., Jarvis, E.D., 2023. Multiple chicken (*Gallus gallus*) genome references to advance genetic variation studies. Single haplotype chicken genome assemblies. *Cytogen. Genome Res.* 162, 405–528. <https://doi.org/10.1159/000529376>.