**Bioinformatic Processing and Analysis:**

The CASAVA (Illumina) program is used for the initial processing of the samples by converting the ".bcl" (base call files) to ".fastq" extensions that are compatible with the programs used for the alignment of reads. Stored FASTQ format files related with each read contain an identification key, a barcode containing the relevant information of the respective sample. The individual barcodes and the remnants of the *PstI* cut site (Figure 2) are the main information used for the demultiplexing of the reads. From the “.fastq file” the expected fragments are categorized into individual files, which correspond to individuals identified by their respective barcodes. The quality of the reads is checked using FastQC v.0.11.3 (Andrews, 2010). Quality trimming is performed in short read sequences during the data process. Both for SNP/CNV calling and for methylation analyses, quality-trimmed reads are aligned against the human reference genome (GRCh38/hg38, UCSC Genome Browser) by selecting the ‘very sensitive-local alignment’ option in the Bowtie2 tool v.2.2.5 pipeline(Langmead and Salzberg, 2012), and using default parameters for paired-end sequences. The *PstI*-reduced genome is use as input and for SNP calling. The input works as genomic background correction in the epigenetic analysis. Samtools v.0.1.19 (Li *et al.*, 2009) is used to evaluate the coverage depth of each sample.

For SNP call, the Tassel v.3.0 program (Glaubitz *et al.*, 2014) is used following the default TASSEL-GBS Discovery Pipeline. Filtering parameters of 1% for minimum minor allele frequency (mnMAF), 70% of minimum taxon call rate (mnTCov), and 70% of site call rate (mnScov) are used for SNP calling on the reduced genomes data. This procedure has been previously described for SNP calling in chickens (Pértille *et al.*, 2016).

Regarding the CNV and methylated DNA sequencing data, the Stacks v.1.39 (Chavez, Jozefczuk, Grimm, Dietrich, Timmermann, Herman, *et al.*, 2010) program is used for data de-multiplexing and quality trimming of reads, with default parameters. DMR calling is performed with data from the methylomic library, while the genomic library is use as the input for genomic background correction and for CNV calling. For CNV calling, the aligned sequence files (.bam) of each individual (from each treatment) are merged into unique files. The “view“ option from Samtools v.1.3.14 (Li *et al.*, 2009) is used to generate a “hit” file from each unique file containing the coverage information for each base pair sequenced in each merged file. This “hit” file is then used for CNV calling by the CNV-Seq too (Li *et al.*, 2009) across the human reference genome using default parameters. Significant CNVs were defined according to the thresholds p' = 10-5 and log2(r') = 0.6, as previously described(Xie and Tammi, 2009). For the DMR calling, the Macs2 (v.2.1.1) (Feng *et al.*, 2012) tool is used after data alignment. Macs2 is a recommended tool to identify sample-wise ‘peak specific’ methylated regions of variable sizes in experiments using paired control to determine enrichment against background (Feng *et al.*, 2012; Niazi *et al.*, 2016; Cavalcante, Qin and Sartor, 2019). This program combines the sequences and defines the peaks (Qvalue=0.1) of sequencing coverage between the cases and their controls, as previously described (Pértille *et al.*, 2020). With this, a file informing the positions of each peak so- called “regions of interest" (ROI) is created, which can be used as input in the downstream analyzes. We employed the BSgenome.Hsapiens.UCSC.hg38 package from the Bioconductor R repository as the reference genome. The MEDIPS R-package was used for basic data processing, quality controls, normalization as well as the recognition of differential methylation regions (DMRs), The MEDIPS package provides a maximum number of stacked reads per genomic position, which reduces possible artifacts caused by PCR amplification. This is achieved by the Poisson genome-wide distribution of the stacked reads. We selected the threshold of P=0.001 as the criterion for the detection of stacked reads. The conversion of the MeDIP data into genome-wide relative methylation scores is performed through a CpG dependent normalization(Chavez, Jozefczuk, Grimm, Dietrich, Timmermann, Lehrach, *et al.*, 2010). A threshold for a minimum sum of counts across all samples per window is defined according to the program default settings (minRowSum=10). Subsequently, the sequencing data of each individual is assigned to one of the experimental groups and differential coverage is calculated between two pre-defined conditions. As our method does not consider CpG´s fragment density to generate relative methylation scores (because our libraries contain fragments with defined start- and end-positions), the MeDIP parameter is set as FALSE. In order to generate the annotation list of the DMRs found, the publicly available genome of interest is interrogated. In the present case, the human reference genome (GRCh38.p12, NCBI) is interrogated with the Variant Effect Predictor (VEP) tool (McLaren *et al.*, 2016).

**References**

Andrews, S. (2010) ‘FASTQC. A quality control tool for high throughput sequence data’, *Reference Source*.

Cavalcante, R. G., Qin, T. and Sartor, M. A. (2019) ‘Novel Bioinformatics Methods for Toxicoepigenetics’, in *Toxicoepigenetics*. Elsevier, pp. 265–288. doi: 10.1016/B978-0-12-812433-8.00012-5.

Chavez, L., Jozefczuk, J., Grimm, C., Dietrich, J., Timmermann, B., Lehrach, H., *et al.* (2010) ‘Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage’, *Genome Research*, 20(10), pp. 1441–1450. doi: 10.1101/gr.110114.110.

Chavez, L., Jozefczuk, J., Grimm, C., Dietrich, J., Timmermann, B., Herman, J. G., *et al.* (2010) ‘Evaluation of single CpG sites as proxies of CpG island methylation states at the genome scale’, *Nucleic Acids Res*. 2013/07/23. Edited by D.-J. De Koning. BioMed Central, 8(7), p. e67500. doi: 10.1371/journal.pone.0067500.

Feng, J. *et al.* (2012) ‘Identifying ChIP-seq enrichment using MACS’, *Nature Protocols*. 2012/07/12. Edited by N. A. Tinker. Cham: Nature Publishing Group, 7(9), pp. 1728–1740. doi: 10.1038/nprot.2012.101.

Glaubitz, J. C. J. *et al.* (2014) ‘TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline.’, *PloS one*. Edited by N. A. Tinker. Public Library of Science, 9(2), p. e90346. doi: 10.1371/journal.pone.0090346.

Langmead, B. and Salzberg, S. L. (2012) ‘Fast gapped-read alignment with Bowtie 2’, *Nature Methods*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 9(4), pp. 357–359. doi: 10.1038/nmeth.1923.

Li, H. *et al.* (2009) ‘The sequence alignment/map format and SAMtools’, *Bioinformatics (Oxford, England)*, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

McLaren, W. *et al.* (2016) ‘The Ensembl Variant Effect Predictor’, *Genome Biol*. 2016/06/09, 17(1), p. 122. doi: 10.1186/s13059-016-0974-4.

Niazi, U. *et al.* (2016) ‘DISMISS: detection of stranded methylation in MeDIP-Seq data’, *BMC Bioinformatics*, 17(1), p. 295. doi: 10.1186/s12859-016-1158-7.

Pértille, F. *et al.* (2016) ‘High-throughput and Cost-effective Chicken Genotyping Using Next-Generation Sequencing’, *Scientific Reports*, 6(January), p. 26929. doi: 10.1038/srep26929.

Pértille, F. *et al.* (2020) ‘Putative Epigenetic Biomarkers of Stress in Red Blood Cells of Chickens Reared Across Different Biomes’, *Frontiers in Genetics*, 11(Llc), pp. 1–44. doi: 10.3389/fgene.2020.508809.

Xie, C. and Tammi, M. T. (2009) ‘CNV-seq, a new method to detect copy number variation using high-throughput sequencing’, *BMC Bioinformatics*, 10(1), p. 80. doi: 10.1186/1471-2105-10-80.