A genome-wide epistatic network underlies the molecular architecture of continuous color variation of body extremities

Julie Demars, Yann Labrune, Nathalie Iannuccelli, Alice Deshayes, Sophie Leroux, Hélène Gilbert, Patrick Aymard, Florence Benitez, Juliette Riquet

Genomics 114 (2022) 110361

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

A genome-wide epistatic network underlies the molecular architecture of continuous color variation of body extremities

Julie Demars [a,*], Yann Labrune [a], Nathalie Iannuccelli [a], Alice Deshayes [b], Sophie Leroux [a], Hélène Gilbert [a], Patrick Aymard [a], Florence Benitez [a], Juliette Riquet [a]

[a] GenPhySE, Université de Toulouse, INRAE, ENVT, Toulouse INP, F-31326 Castanet-Tolosan, France
[b] UMR967, CEA, INSERM, Institut de Radiobiologie Cellulaire et Moléculaire, Télomères et réparation du chromosome, F- 92265 Fontenay-aux-Roses, France

ARTICLE INFO

Keywords:
Coat coloration
Copy number variation, epistasis
Melanogenesis
Recessivity

ABSTRACT

Deciphering the molecular architecture of coat coloration for a better understanding of the biological mechanisms underlying pigmentation still remains a challenge. We took advantage of a rabbit French experimental population in which both a pattern and a gradient of coloration from white to brown segregated within the himalayan phenotype. The whole experimental design was genotyped using the high density Affymetrix® AxiomOrcun™ SNP Array and phenotyped into 6 different groups ordered from the lighter to the darker. Genome-wide association analyses pinpointed an oligogenic determinism, under recessive and additive inheritance, involving genes already known in melanogenesis (*ASIP*, *KIT*, *MC1R*, *TYR*), and likely processed pseudogenes linked to ribosomal function, *RPS20* and *RPS14*. We also identified *(i)* gene-gene interactions through *ASIP:MC1R* affecting light cream/beige phenotypes while *KIT*:*RPS* responsible of dark chocolate/brown colors and *(ii)* a genome-wide epistatic network involving several others coloration genes such as *POT1* or *HPS5*. Finally, we determined the recessive inheritance of the English spotting phenotype likely involving a copy number variation affecting at least the end of the coding sequence of the *KIT* gene. Our analyses of coloration as a continuous trait allowed us to go beyond much of the established knowledge through the detection of additional genes and gene-gene interactions that may contribute to the molecular architecture of the coloration phenotype.

1. Introduction

Understanding the molecular mechanism of coloration has been the goal of many genetic and evolutionary studies in a broad number of species [1–3]. Specific color-producing cells contribute to animal coloration and patterns. The so-called "dermal chromotophore unit" [4] involves several types of chromatophores, including pterin, carotenoid, iridophores, and melanophores or melanocytes producing chemically melanin pigments. More than a hundred genes have been involved in coloration traits in model species such as drosophila or mice but also in wild species [5–8]. In mammals, phenotypic characteristics of animal coloration may be classified based on the patterning and/or type and amount of pigment produced through melanogenesis pathway [9]. The main genes that alter the development of melanocytes are the Proto-Oncogene Receptor Tyrosine Kinase (*KIT*), the Microphthalmia associated Transcription Factor (*MITF*) and the endothelin axis [10,11]. Other genes such as Tyrosinase (*TYR*), Tyrosinase related Protein 1 (*TYRP1*), the Oculocutaneous albinism 2 (*OCA2*) and the Membrane-Associated Transporter Protein (*MATP*) affect melanin synthesis. Another group of genes related to pigment synthesis are those that control the switch between eumelanin and pheomelanin production. Those with the strongest effect in this change are the Melanocortin 1 Receptor (*MC1R*) and the Agouti Signalling Protein (*ASIP*).

While many studies considered color traits as complex phenotypes analysing skin or hair colors as categories [12–14], most of recent analyses evaluated pigmentation as continuous variations [15–18]. Although GWAS have allowed for a greater understanding of the genetic component of many complex traits, the genetic effects highlighted are

Abbreviations: BAF, B allele frequency; CNV, copy number variation; CS, credible set; GWAS, genome wide association study; LD, linkage disequilibrium; LFSR, local false sign rate; LRR, log R ratio; Ocu, *Oryctolagus cuniculus*; PIP, posterior inclusion probability; SNP, single nucleotide polymorphism.

* Corresponding author.
E-mail addresses: julie.demars@inrae.fr (J. Demars), yann.labrune@inrae.fr (Y. Labrune), nathalie.iannuccelli@inrae.fr (N. Iannuccelli), alice.deshayes@cea.fr (A. Deshayes), sophie.leroux1@inrae.fr (S. Leroux), helene.gilbert@inrae.fr (H. Gilbert), patrick.aymard@inrae.fr (P. Aymard), florence.benitez@inrae.fr (F. Benitez), juliette.riquet@inrae.fr (J. Riquet).

https://doi.org/10.1016/j.ygeno.2022.110361
Received 17 November 2021; Received in revised form 22 March 2022; Accepted 29 March 2022
Available online 1 April 2022

Table 1
Genes and mutations responsible of coat coloration phenotypes in rabbits.

Locus	Gene	Location	Allele	Mutation	Phenotype	Reference
A (Agouti)	*ASIP*	Ocu4	$A > a^t > a$	A = wt a^t = 11Kb deletion a = c.5_6insA	agouti tan non-agouti	27,60
B (Brown)	*TYRP1*	Ocu1	B > b	B = wt b = g.41360196	black brown	25
C (Color)	*TYR*	Ocu1	$C > C^{ch} > C^h > c$	C = wt C^{ch} = T358I C^h = E294G c = c.304_333del30	full color chinchilla himalayan albino	26
D (Dilution)	*MLPH*	GL018840	D > d	D = wt d = g.549853delG	non-dilute dilute	24
E (Extension)	*MC1R*	GL018965	$E^d > E^s > E > e^j > e$	E^d = c.2806285del6 E^s = unknown E = wt e^j = c.[124A;125_130del6] e = c304_333del30	black steel full extension japanese non-extension	61,62
En (English Spotting)	*KIT*	Ocu15	En > en	En = unknown en = wt	spotted non-spotted	55

Genomic locations correspond to either chromosome, (Ocu1 for *O. cuniculus* chromosome 1) or scaffold (GL018840) based on the European rabbit reference genome OryCun2.0.

largely small and often focused on common SNP and additive genetic models. More and more studies explore alternative heritable components such as genetic interactions but there are still some challenges for identifying significant epistasis [19]. Skin, hair/coat pigmentation represent then pertinent phenotypes since the genetic determinism of coloration determined is likely polygenic involving a few genes with large effects [14,20]. Typically, in the case of coat coloration, molecular interactions are known between MC1R and its antagonist ASIP peptide since gain-of-function *ASIP* mutations block MC1R signalling and lead to the production of red pheomelanin [21,22]. Epistatic interaction of those two genes modulates wool color in creole sheep breed [74]. Moreover, strong synergistic interactions have also been highlighted for other color traits such as skin/hair pigmentation in humans for which an interaction between *HERC2* and *MC1R* has been shown to be significantly associated [20].

In the European rabbit (*Oryctolagus cuniculus*), different coat colors have been selected through domestication and are nowadays fixed in specific breeds. Therefore, candidate gene approaches have allowed the identification of various mutations responsible of different phenotypes such as the dilution of the coat color [24] or the brown phenotype [25]. In rabbits, six loci (called A for Agouti, B for Brown, C for Color, D for Dilution, E for Extension and En for English Spotting) are involved in the coloration of the coat (Table 1). Notably, allelic heterogeneity with dominance/recessivity relations exists between different mutations within the single loci, resulting in distinct phenotypes. For instance, at the C locus (*TYR*), the C^{ch} mutation, responsible of the chinchilla phenotype, is dominant over the C^h mutation, itself associated with himalayan coat coloration which is dominant over the *c* mutation leading to albino phenotype [26] (Table 1). In addition, epistatic effects between both Extension and Agouti loci have been shown from a cross between a Champagne d'Argent buck and a Thuringian doe [27].

Although the mutation responsible of the himalayan phenotype has been identified, this trait has always been described in a simple way without considering both the gradient and the pattern of coloration occurring within the phenotype. A better insight of the coloration variability requires a fine characterization of the phenotype to highlight dominance and/or recessive effects and epistatic interactions. Here, we propose a genome-wide investigation of coat color of body extremities using the high-density SNP rabbit beadchip (Affymetrix® AxiomOrcun™ SNP Array) in an experimental familial design. We *(i)* identified several significant loci including key genes involved in melanogenesis but also candidate genes which are processed pseudogenes linked to ribosomal proteins, *(ii)* highlighted how epistatic phenomena likely contribute to the genetic determinism of color variation of body extremities and *(iii)* determined the recessive inheritance of the English spotting phenotype likely involving a copy number variation within the *KIT* gene.

2. Material and methods

2.1. Animal data

2.1.1. Ethics approval

The French ministery of higher education, Research and innovation and the local animal research ethics committee (C2EA-115) approved the study (approval number 00903.01). All procedures were conducted in accordance with the French legislation on animal experimentation and ethics. JD was authorized by the French Ministry of Agriculture to conduct experiments on living animals (approval number 312011116).

2.1.2. The experimental design

The experimental rabbit populations were issued from the INRA 1001 line [28] and bred in the INRAE experimental farm in accordance with the national regulations for animal care. The experimental population was a combination of two genetically related lines: the G10 line, selected for 10 generations for decreasing Residual Feed Intake (RFI) [29] and the G0 control line produced from frozen embryos of the ancestral population of the selected line. The 296 G10 and 292 G0 rabbits were produced in the same 3 batches with a 42 days interval. In each batch, half of the kits was fostered to G0 does and the second half was fostered by G10 does. Adoptions of offspring were managed alternatively between both lines in successive batches. At weaning (32 days), in each batch, kits were placed in individual cages. More details about the experimental cross can be found in Garreau et al. [30] Briefly, the initial design included 832 rabbits including the 20 bucks, 101 does and 711 offspring. Although the experimental cross was not designed for evaluating coloration traits, we took advantage of it since the G0 line originated from Californian rabbits. We observed within the experimental cross a segregation for both color and pattern variability within the himalayan phenotype. The final design, based on phenotypic evaluation is detailed further and on Supplementary Fig. S1.

2.1.3. Phenotypic data and quality control

We first distinguished 5 different rabbit color groups from white to dark by visual inspection of the whole population. This notation was performed by 2 independent experimenters. Colors were classified as A, B, C, D and E. Secondly, we selected a few individuals ($n = 15$ per group) that were phenotyped for their nose coloration using a colorimeter to

validate our subjective classification. A significant correlation was observed between the luminescence (L*) measurement and the notes (Supplementary Fig. S1a), validating the determined groups. Moreover, an additional group was created since some animals from class A had colored ears but white noses. Altogether, 6 ordered phenotypes were defined, sorted from P1 to P6 and numbered 1 to 6 for further quantitative analyses (P1 = 1, P2 = 2, P3 = 3, P4 = 4, P5 = 5, P6 = 6). In total, 686 rabbits out of the 832 of the whole experimental design were assigned to one phenotypic group, including 574 offspring, 20 bucks and 92 does, with 2 to 50 offspring per buck and 1 to 15 offspring per doe (Supplementary Fig. S1b). Number of rabbits per phenotype were 34, 32, 47, 125, 193 and 255 for P1, P2, P3, P4, P5 and P6, respectively (Supplementary Fig. S1c). Classical plots were built with the ggplot2 R package (https://www.rdocumentation.org/packages/ggplot2) [31].

2.1.4. Sampling collection and DNA extraction

Ear punch biopsies were collected in Allflex Tissue Sampling Unit tube (Allflex France, Vitré, France) and genomic DNA was extracted from samples with a home-made protocole: proteinase K lysis following by salt-based DNA extraction and ethanol precipitation. Briefly, ear punch biopsies were digested at 56 °C for 3 h using a 500 µL solution including 10 mM Tris HCl, 0.1 M EDTA pH = 8, 0,5% SDS and 0.2 mg proteinase K. After overnight incubation at 37 °C, 1/3 volume of saturated (6 M) NaCl were added and slightly mixed before a centrifugation step (30 min at 4 °C and 21,000 g). The supernatant was mixed with 2 volumes of 100% ethanol. DNA was retrieved and resuspended in classic buffer for 1 h at 60 °C before an overnight resuspension at 37 °C. Total genomic DNA was quantified using the Nanodrop 8000 (ND8000LAPTOP, Thermo Fisher Scientific, USA) and the Qubit2.0 (Q32866, Life Technologies, USA).

2.1.5. Genotyping data and quality control from SNP array

The DNA samples were genotyped at the Centro Nacional de Genotipado (CeGen) platform (Santiago de Compostela, Spain) using the Affymetrix® AxiomOrcun™ SNP Array as recommended by the manufacturer. The SNP array contains 199,692 molecular markers spanning both chromosomes and scaffolds. The order of the SNPs was based on the Rabbit OryCun2.0 assembly released by the Broad Institute of MIT and Harvard [32]. Missing data imputation and haplotype phasing were performed with the software FImpute [33]. The SNP data were then filtered based on minor allele frequencies ≥0.005 leading to a final SNP dataset of 162,070 markers for association analyses. In a first approach, we did not apply standard filtering parameters including call rate, call frequency and Hardy-Weinberg disequilibrium that may lead to the elimination of SNPs that overlapped structural variants. The observation of significant deviation from Hardy-Weinberg equilibrium is the first clue of the presence of a structural variant [34]. In addition, those SNPs may not form three discrete clusters when plotting allele-A intensity *vs.* allele-B intensity in the genotyping software analysis [35]. However, the quality (call rate (95%), call frequency (95%), Hardy-Weinberg disequilibrium (10^{-06})) of highlighted variants lying within intervals of interest were checked *a posteriori* to secure our results especially for epistasis analyses. Additional 12,640 SNP were excluded for genome-wise epistasis study.

2.1.6. Manual genotyping data

Manual genotyping of 5 variants, included 4 known mutations (*ASIP* - allele a, *TYR* - alleles c and C^h and *MC1R* - allele Ed) and another variant for ASIP - 5,435,370 bp were performed. SNPs were genotyped either using RFLP PCR (TYR - alleles c and C^h) or allele-specific PCR (ASIP - 5,435,370 bp). Deletions were genotyped using Capillary Electrophoresis (*ASIP* - allele a and *MC1R* - allele Ed and e). Primers and PCR conditions used are presented in Supplementary Table S1.

For RFLP PCR, PCR were performed with the kit GoTaq® Flexi (Promega, USA) using 20 ng DNA, 0.5 mM of primers, 0.2 mM dNTPs (Promega, USA), 1× buffer, 1.5 mM MgCl2 et 0.25 U Taq in a final volume of 12 µL. Digestions were performed with *Nci*I I et *Bsa*XI (NEB, USA) for TYR - alleles c and C^h, respectively, using 2 U of enzyme and 1× of their respective buffer before incubation at 37 °C for 15 min. The PCR and digestion were performed on thermocycleur Verity (Thermo Fisher Scientific, USA) and PCR products were loaded on a 2.5% agarose gel with ethidium bromide.

For allele-specific PCR, we used the KASPAR (Kompetitive Allele Specific PCR) (KBioscience, United Kingdom) technology. Amplification was performed with 10 ng DNA, 1× PCR buffer, 1.8 mM MgCl2, 0.2 mM dNTPs, 0.25 µM of each fluorescent dye (Fam et Vic), 0.5 U of Taq polymerase and 12 µM for allele-specific primers and 30 µM for the common primer in a final volume of 5 µL. We followed provider recommendation for the PCR program, fluorescent reading was made on a Quant Studio 6 (Thermo Fisher Scientific, USA) and results were analysed with the software Quant studio Real Time PCR (Thermo Fisher Scientific, USA).

Genotyping using capillary electrophoresis were performed on a ABI3730™ (Applied Biosystems, USA). The PCR were performed with 0.1 mM of the extended primer, 0.15 mM of the hybridization primer carrying the dye and 0.15 mM of the reverse primer. The other conditions of PCR and cycle are similar to the RFLP PCR. PCR products were loaded on the ABI3730™ after a first step of 1/20 dilution and 2 µL of the dilution were mixed with formamide and size standard GeneScan-600Liz Size Standard (Applied Biosystems, USA) before a denaturation step at 94 °C for 5 min. Analyses were perfomed with GeneMapper™ Software (Applied Biosystems, USA).

2.1.7. RNA-seq alignments

Publicly available RNA-seq raw data from back skin of rabbits were uploaded to perform alignment, quantification and transcript discovery with statistics. Three Rex rabbits with black or white or chinchilla back skin were considered. Accession numbers for BioSample (https://www.ncbi.nlm.nih.gov/biosample/) are SAMN02693835, SAMN02693836 and SAMN02693834 for black, white and chinchilla, respectively. Accession number for raw data (https://trace.ncbi.nlm.nih.gov/Traces/sra/) are SRR1201255, SRR1201256 and SRR1201257 for black, white and chinchilla, respectively. Quality controls, alignments and analyses were performed with the open-source nf-core/rnaseq workflow (https://nf-co.re/rnaseq) using the 3.0 version that implemented fastqc 0.11.9 and qualimap 2.2.2 for quality controls and STAR 2.6.1 for the mapping. Paired-end reads were aligned on the reference OryCun2.0 genome and annotation version Oryctolagus_cuniculus.OryCun2.0.104.gtf was used. Mapping of RNA-seq experiments were showed using the Integrative Genome Viewer (https://software.broadinstitute.org/software/igv/) [36].

2.2. Statistical analyses

2.2.1. Univariate linear mixed models for association analyses

We used the GEMMA (Genome-wide Efficient Mixed Model Association) software to perform association analyses. Briefly, GEMMA fits a univariate linear mixed model (LMM) [37] or a Bayesian sparse linear mixed model using Markov chain Monte Carlo (BSLMM) [38]; both methods control for population structure.

SNP effects were tested with the following univariate animal mixed model LMM [37]:

$$\boldsymbol{y} = \boldsymbol{W\alpha} + \boldsymbol{x}\beta + \boldsymbol{u} + \boldsymbol{\epsilon} \text{ with } \boldsymbol{u} \sim \mathrm{MVN}_n\left(0, \lambda\tau^{-1}\mathbf{K}\right) \text{ and } \boldsymbol{\epsilon} \sim \mathrm{MVN}_n\left(0, \tau^{-1}\mathbf{I_n}\right)$$

where $\boldsymbol{y}$ is the vector of phenotypes for a given trait, $\boldsymbol{W}$ is the incidence matrix of covariates corresponding to fixed effects and $\boldsymbol{\alpha}$ stands for the effects of these covariates, $\boldsymbol{x}$ is the vector of allelic dosages of the genotypes (0, 1 or 2) and β stands for marker size effect, $\boldsymbol{u}$ is the random polygenic effect and $\boldsymbol{\epsilon}$ is the random residual effect. Additive effects are structured after K, the centered relatedness matrix computed from the genotypes, λ is the ratio between the two variance components and $\boldsymbol{\tau}$ is

the variance of the residual errors. In addition to the classical additive model, dominant and recessive models were tested by transforming allelic dosages of the genotypes into binary genotypes based on the minor allele at each SNP. Association was concluded as *(i)* significant at genome-wise level after a Bonferroni correction ($0.05 / 162{,}070 = 3.08 \times 10^{-07}$) and *(ii)* suggestive at chromosome-wise level after a Bonferroni correction (0.05 / n markers on Ocu). Moreover, we used the genotypes from manual genotyping of known mutations as covariates in the model to show whether these mutations might explained significant association signals.

The SNP effects were also tested with the following Bayesian sparse animal mixed model [38]:

$$y = 1_n\mu + \boldsymbol{X}\beta + \boldsymbol{u} + \boldsymbol{\epsilon}$$
$$\beta_i \sim \pi \mathrm{N}\left(0, \sigma_a^2\tau^{-1}\right) + (1-\pi)\delta_0, \boldsymbol{u} \sim \mathrm{MVNn}\left(0, \sigma_b^2\tau^{-1}\mathbf{K}\right), \boldsymbol{\epsilon} \sim \mathrm{MVNn}\left(0, \tau^{-1}\mathbf{I}_n\right)$$

where $\mathbf{1_n}$ is a n-vector of 1 s, μ is a scalar representing the phenotype mean, X is an n x p matrix of genotypes measured on n individuals at p genetic markers, β is the corresponding p-vector of the genetic marker effects, and other parameters are the same as defined in the standard linear mixed model. In the special case $K = XX^T/p$, the SNP effect sizes can be decomposed into two parts: α that captures the small effects that all SNPs have, and β that captures the additional effects of some large effect SNPs. In this case, $\boldsymbol{u} = X\alpha$ can be viewed as the combined effect of all small effects, and the total effect size for a given SNP i is $\alpha_i + \beta_i$. To pinpoint signals, we summed the sparse probabilities evaluated from the total effect size for a given SNP on sliding windows containing 20 SNPs.

Manhattan and Quantile-Quantile plots were built using the qqman R package (https://www.rdocumentation.org/packages/qqman) [39].

2.2.2. Fine-mapping of regions of interest

We also used the SuSiE (Sum of Single Effects) model, which corresponds to a new formulation of the Bayesian variable selection in regression (BVSR) to fine-map the loci [40]. This model fits an Iterative Bayesian Stepwise Selection (IBSS) algorithm that is a Bayesian analogue of traditional stepwise selection methods. SuSie produces Posterior Inclusion Probabilities (PIPs) and Bayesian Credible Sets (CSs) which capture an effect variable allowing the fine-mapping of significant detected regions. The SuSiE method removes the single causal variant assumption and groups SNPs into distinct association signals in the analysis, such that it aims to find as many CSs of variants that are required so that each set captures an effect variant, whilst also containing as few variants as possible. Regions showing either $-\log_{10}(p\text{-value}) > 6.5$ with the LMM (corresponding to 5% genome-wide threshold after a Bonferonni correction) or a PIP > 0.1 were consider for further analyses.

2.2.3. Epistatic interaction analyses

A first evaluation of epistasis was performed using the linear regression model implemented in PLINK 1.9 [41] with the –epistasis option to fit the model:

$$y = \beta_0 + \beta_1 g_A + \beta_2 g_B + \beta_3 g_A g_B$$

for each inspected variant pair (A, B), where g_A and g_B are allele counts, and $g_A g_B$ is the count of common occurrences of the alleles at the two loci; then the β coefficients are tested for deviation from zero. Pairwise interaction was tested between each marker of the set of the 7 best associated SNP variants. Interactions with p-value <0.05 are considered significant.

To determine the best classification of individuals within the 6 phenotypig groups (P1 to P6) given their combined genotypes at the different selected regions, we built a decision tree using the Classification And Regression Trees (CART) algorithm [83]. Moreover, to identify the most likely epistatic interactions between the selected markers, we used the decision criterion BIC metric. The stepAIC function in the MASS package was applied [43]. Decision trees were built with the rpart R package (https://www.rdocumentation.org/packages/rpart) [44]; the Fig. 3b is a concatenation of two independent decision trees.

Pairwise epistatic interactions between the set of selected markers from GWAS and the rest of the genome were also evaluated using the adaptive shrinkage method [45] implemented in the ashr R package (https://www.rdocumentation.org/packages/ashr/versions/2.2-47). Both interaction effect size and corresponding standard error were thus estimated for each pairwise combination. Those measures were used with an empirical Bayes approach for large-scale hypothesis testing. This method accounts for variation in measurement precision across tests in the computation of the effect sizes, and facilitates their estimation. In addition, instead of p-value, q-value or local FDR, the "Local False Sign Rate" (lfsr) [45], which refers to the probability of getting the sign of an effect wrong, was computed.:

$$lfsr_j := min\left[\, Pr\left(\beta_j \geq 0 | \widehat{\pi}, \widehat{\beta}, s\right), Pr\left(\beta j \leq 0 | \widehat{\pi}, \widehat{\beta}, s\right)\right]$$

with for effect j, lfsr_j, is the probability that we would make an error in the sign of effect β_j if we were forced to declare it either positive or negative. Small values of $\mathrm{lfsr_j}$ indicate that we can be confident in the sign of β_j, which implies that we are confident it has a non-zero value. Lfsr is a more conservative measure of significance than local FDR and it is more robust to modeling assumptions. The circular plot was built using the BioCircos R package (https://www.rdocumentation.org/packages/BioCircos) [46].

2.2.4. Linkage disequilibrium pattern of intervals

Pairwise linkage disequilibrium measure (r [2]) was computed using the PLINK 1.9 software with the –r2 option (www.cog-genomics.org/plink/1.9/)[41] for all the pairs of SNPs of the selected regions. Linkage disequilibrium profiles were visualized with the LDheatmap R package (https://www.rdocumentation.org/packages/LDheatmap) [47].

2.2.5. Linkage map construction

The netmap option of the netgwas R package [48] was used for building a linkage map of the Ocu14 region since 2 distinct intervals away from 20 Mb were within the same credible set. An outbred population, such as full-sib families here, derived from mating two non-homozygous parents results in markers in the genome of progenies that cannot be easily assigned to their parental homologues. Neighboring markers that vary only on different haploids will appear as independent, therefore requiring a different ordering algorithm. In that case, the reverse Cuthill-McKee (RCM) algorithm [49] was used to order markers. The RCM algorithm is based on graph models.

2.2.6. Copy number variations evaluation

Genotype array data provides measurements at each of hundreds of thousands of genetically variable sites, allowing discovery of genome structural variation besides of genotyping [28]. The raw measurements consist of two intensity signals, one for each allele, which are subsequently transformed into the log-scaled ratio of the observed and the expected intensity (LRR), and the B Allele Frequency (BAF) which captures the relative contribution from one allele (B) to the fluorescent signal. While expected values of 0 for LRR reflect normal copy number ($n = 2$ for diploid individuals, log2(2/2)), aberrant theoretical values of 0.57 or − 1 reflect one copy gain (log2(3/2)) or loss (log2(1/2)), respectively. From BAF values, a BAF value of 0.5 indicates a heterozygous genotype (AB), whereas 0 and 1 indicate homozygous genotypes (AA and BB, respectively). For example, a single copy number gain is characterised by 4 theoretical distinct BAF values = 0, 0.33, 0.67 and 1, reflecting AA, AAB, ABB and BB genotypes, respectively. Instead of the characterization of individual structural variants, we analysed abnormal values of LRR and BAF in the different phenotypic groups. We considered the mean value of LRR at each SNP of the interval in the various extremities-colored rabbits to likely pinpoint CNV shared across

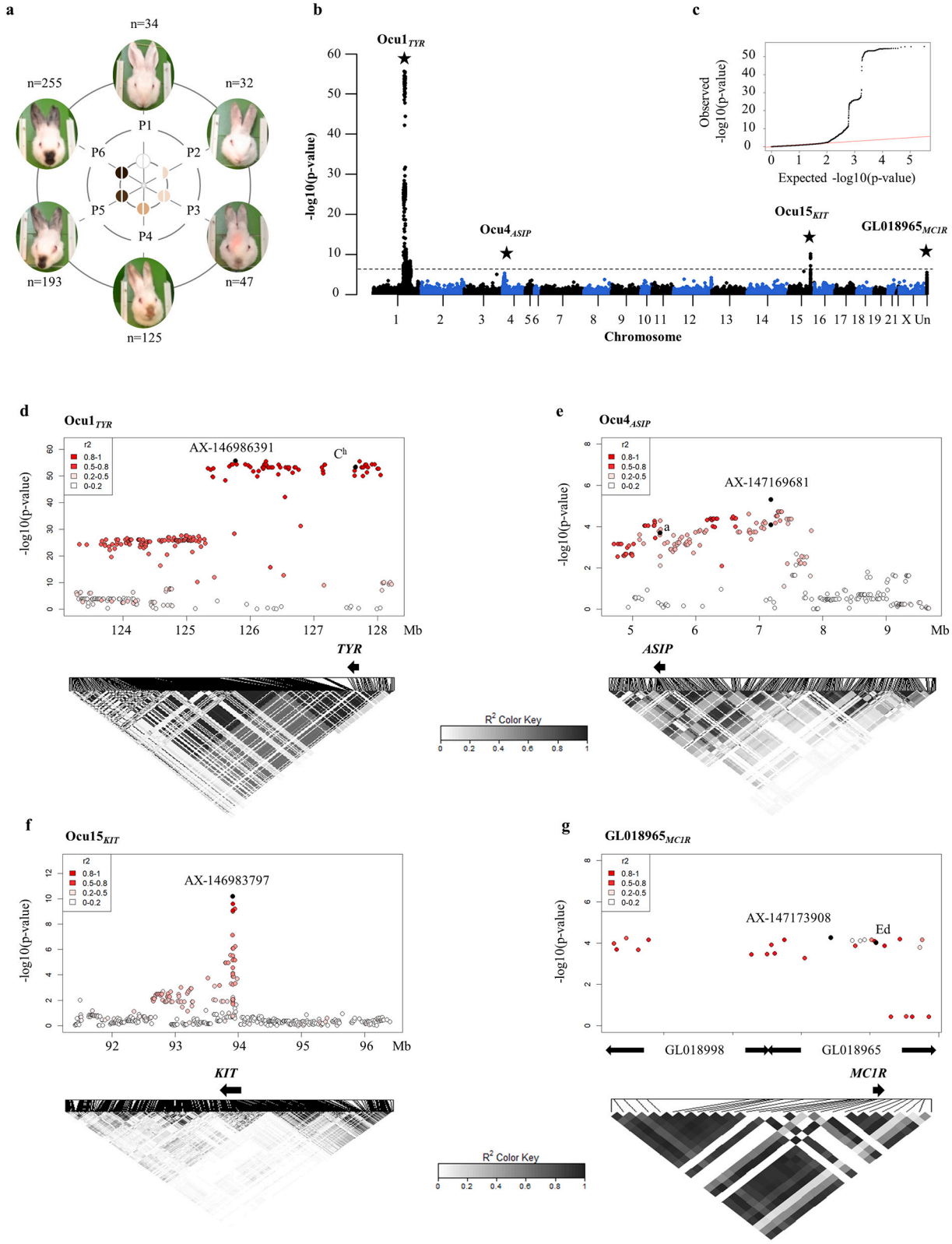

Fig. 1. Coat coloration association results for the whole experimental design. a. Phenotypic classification. All rabbits were phenotyped in order to give them a phenotypic value. In total, 6 distinct coat pigmentations (P1 to P6) were considered and ordered in a gradient ranging from white to light and dark as quantitative phenotypes from 1 to 6. b. Manhattan plot. GWAS was performed using quantitative 1 to 6 phenotypes under a linear mixed model. Location of SNPs on the x-axis is based on the reference OryCun2.0 genome and all scaffolds were regrouped under an extra chromosome Unknown. The dashed line represents the 5% genome-wide threshold. c. Q-Q plot corresponding to the GWAS analysis. Highlights on a few regions of interest d for *TYR* locus on Ocu1 ($Ocu1_{TYR}$), e for *ASIP* locus on Ocu4 ($Ocu4_{ASIP}$), f for *KIT* locus on Ocu15 ($Ocu15_{KIT}$) and g for *MC1R* locus on scaffold GL01865 ($GL018965_{MC1R}$)) with regional Manhattan plots showing the best associated marker and already known mutations, and local linkage disequilibrium heatmap.

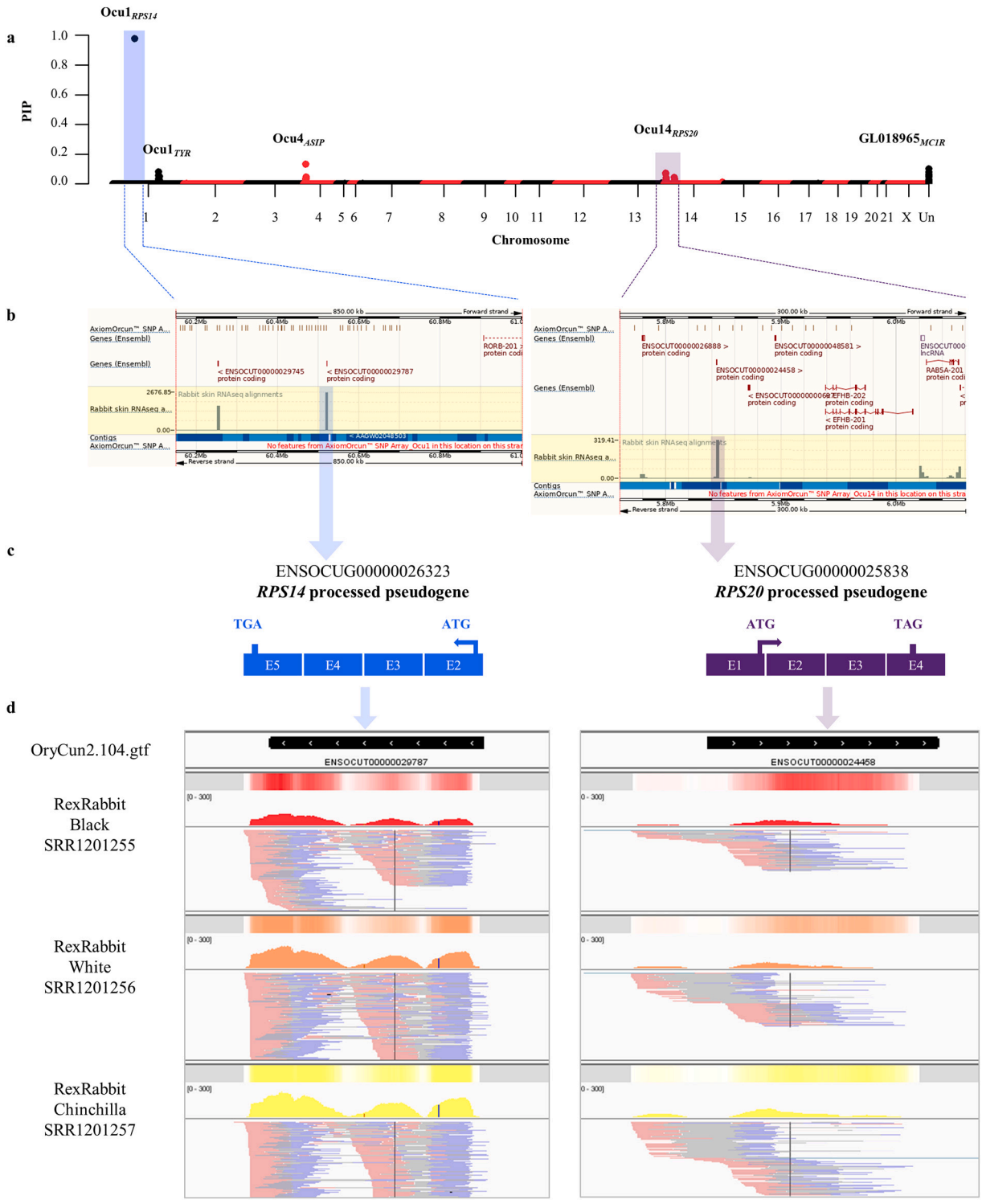

Fig. 2. Highlights on processed pseudogenes of ribosomal proteins (*RPS14* and *RPS20*). a. Manhattan plot depicting posterior inclusion probabilities (PIP), which pinpoint fine mapping of genomic regions involved in cream to dark brown color variability including P3 to P6 colored rabbits. Location of SNPs on the x-axis is based on the reference OryCun2.0 genome and all scaffolds were regrouped under an extra chromosome Unknown. b. Screenshots of both Ocu1$_{RPS14}$ and Ocu14$_{RPS20}$ Ensembl regions. The first track represents the position of SNP on the Affymetrix® AxiomOrcun™ SNP Array, the second track represents the annotated genes and the third track highlighted in yellow shows the quantification of transcripts in the generic rabbit sample (accession number SAMN00013655). c. Representation of both *RPS14* and *RPS20* processed pseudogenes in the OryCun2.0 reference genome. d. Screenview of Intregrative Genome Viewer showing alignments of reads obtained after RNA-seq experiments of skin samples from 3 Rex rabbits carrying black coat (accession number SRR1201255), white coat (accession number SRR1201256) or chinchilla coat (accession number SRR1201255). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

individuals harbouring the same phenotype. A deviation from the expected LRR and BAF normal values for biallelic genotypes was a clue for CNV detection.

Log R Ratio (LRR) and B Allele Frequency (BAF) were extracted from the Axiom™ Analysis Suite Software 4.0.3.3 (Thermo Fisher Scientific, USA) using the Axiom® CNV Summary Tools 1.1 (Thermo Fisher Scientific, USA) after a global analysis of the whole experimental design. The 8 96-well genotyping plates were analysed simultaneously for an accurate definition of genotypes clusters since no reference cluster exist for the Affymetrix® AxiomOrcun™ SNP Array. However, a large difference in signal intensities between plates for both LRR and BAF values were observed for 2 of them likely affecting the results for CNV analyses. The LRR values were normalised after confirming that it was not a biological effect since the 6 phenotypes and all families were represented on these 2 plates. Results presented here are adjusted LRR values taking the plate effect into account using the lm function in R.

3. Results

3.1. Several loci are significantly associated with coat color of body extremities

Coat color of body extremities was analysed as a quantitative trait with phenotypes numbered from 1 to 6 (called P1 to P6) and ordered from lighter to darker (Fig. 1a). A first exploration of the genetic determinism of coat color of body extremities was performed using a simple linear mixed model. A group of more than one hundred markers located on chromosome 1 (called Ocu1 for *O. cuniculus* 1) showed significant associations, with the best signal for the SNP AX-146986391 (p-value $= 2.36 \times 10^{-56}$) (Fig. 1b and Supplementary Table S2). Additional significant and suggestive signals located on Ocu3 (AX-147059932, p-value $= 9.70 \times 10^{-06}$), Ocu4 (AX-147169681, p-value $= 4.84 \times 10^{-06}$), and Ocu15 (AX-146983797, p-value $= 6.80 \times 10^{-11}$) were obtained (Fig. 1b and Supplementary Table S2). In addition, groups of variants located on scaffolds GL018754 (AX-147179313, p-value $= 2.83 \times 10^{-06}$) and GL018965 (AX-147173908, p-value $= 5.49 \times 10^{-05}$), here regrouped for convenience in chromosome Unknown (an arbitrary chromosome that groups together all the scaffolds), also showed a suggestive association (Fig. 1b and Supplementary Table S2).

3.1.1. Two major genes with recessive effects are associated with white and spotting traits

We then focused on the best associated markers to decipher how detected genomic regions contribute to the different coat color of body extremities. Analysis of the genotypic classes for variant $Ocu1_{AX\text{-}146,986,391}$ among the different phenotypic groups showed that all individuals homozygous for the minor allele were P1 animals. Among the P1 class, only 5% of the individuals carry an allele 2, in a heterozygous manner suggesting that this locus, segregated with a recessive inheritance pattern (Supplementary Fig. S2a). To test this assumption, we performed association analyses comparing phenotype 1 (P1) *versus* the 5 remaining ones under different genetic models. A unique significant signal on Ocu1 with the best signal using the recessive model (AX-147087415, p-value $= 1.30 \times 10^{-300}$) was highlighted (Supplementary Fig. S2b). The *TYR* gene is 2 Mb downstream the best associated marker in a region showing high linkage disequilibrium (LD) (Fig. 1d). Already known mutations located in the *TYR* gene, c and C^h, responsible of albino and himalayan phenotypes, respectively, did not show highest significant signals (Fig. 1d and Supplementary Table S3).

In a similar way, an excess of homozygote for the minor allele was observed for the best associated marker located on Ocu15 (Supplementary Fig. S3a), suggesting that the light spotted color corresponding to phenotype P2 is mainly a Mendelian trait. We carried out a GWAS comparing phenotype 2 (P2) to combined light to dark brown extremities-colored rabbits (P3 to P6), excluding phenotype 1 (P1). The association signals for P2 was explained by variants of the chromosome 15 and scaffold GL018754 with the best p-values under a recessive model ($Ocu15_{AX\text{-}146,983,797}$, p-value $= 1.26 \times 10^{-144}$ and $GL018754_{AX\text{-}147,115,616}$, p-value $= 1.72 \times 10^{-37}$) (Supplementary Fig. S3b). The best associated marker on Ocu15 is located within the *KIT* gene (Fig. 1f). For the AX-147115616 SNP from scaffold GL018754, it is close to the *GSX2* and *PDGFRA* genes, which are neighbours to the *KIT* gene in many species. This suggests that GL018754 is likely linked to Ocu15 as also suggested on the LD heatmap (Supplementary Fig. S4), and only one signal should be considered.

3.1.2. Five additional loci account for the remaining coat color of body extremities

To better decipher the molecular architecture of the other four phenotypic groups, we only considered those individuals ($n = 620$) in further analyses. We used a Bayesian sparse linear mixed model, a more appropriate method for polygenic traits. An estimation of 5 to 7 QTLs contributed to coloration phenotypes 3 to 6 (Supplementary Fig. S5a). While 50% of the variance in phenotypes was explained by this model (Supplementary Fig. S5b), most of the genetic variance seemed due to QTLs (Supplementary Fig. S5c). We summed the sparse probabilities for the SNP inclusion on sliding windows containing 20 SNPs to amplify the identified signals from single variants (Supplementary Fig. S5d). Two loci with large effects of approximately 0.6, located on Ocu1 and scaffold GL018965, showed high probabilities of being QTLs (70% and 78%, respectively). The *MC1R* gene belongs to the scaffold GL018965. Some known mutations within *MC1R* did not segregate in this population (japanese and extension alleles), but the best signal (AX-147194100) on GL018965 corresponded to the black dominant E^d mutation of the *MC1R* gene (Fig. 1g and Supplementary Table S3). Concerning the novel position on Ocu1 (AX-146995791), it matches to a gene-poor region on the OryCun2.0 genome assembly. Two additional QTLs, located on Ocu1 and Ocu4, showed intermediate probabilities (36% and 55%, respectively), but also large effects of 0.9 and 0.4, respectively. The highlighted region on Ocu1 spanned the *TYR* locus and signal located on Ocu4 is approximately 1.5 Mb downstream the *ASIP* gene with a long structure of LD as previously shown for Ocu1 (Fig. 1e). Although several variants within *ASIP,* including the agouti a marker, were genotyped, they did not have the best p-values (Fig. 1e and Supplementary Table S3). Finally, 2 novel QTLs, located on Ocu13 and Ocu14, showed a trend of being QTLs with a probability above 15% and a large effect of 0.7 for the QTL on Ocu14.

We fine-mapped identified regions in minimal Credible Set (CS) and validated 5 out of the 6 identified regions (exception of the Ocu13 locus) (Fig. 2a and Supplementary Table S4). One CS contained one SNP (AX-146995791) and the 4 others contained between 16 and 61 markers, with interval sizes ranging from 430 Kb to 2.2 Mb. The scaffold GL018965 was highlighted within a CS including also the scaffold GL018998 (Supplementary Table S4). While the region on GL018965 contained the Extension locus characterised by the *MC1R* gene, the interval on GL018998 pinpointed towards the *ANKRD11* gene. Homologous regions in human or mice highlighted the *MC1R* gene 600 Kb downstream the *ANKRD11* gene confirming that both scaffolds GL018998 and GL018965 might be linked, as confirmed by the LD heatmap (Fig. 1g). A unique CS of markers was identified on Ocu14 regrouping 2 groups of markers located more than 25 Mb away on the chromosome (Supplementary Table S4). Although these two groups of markers were located apart on the draft, linkage analysis using the familial meiosis of our pedigree indicated that these two groups were linked (Supplementary Table S5).

To likely identify candidate genes belonging to novel highlighted Ocu1 and Ocu14 intervals, we analysed publicly available RNA-seq data extracted from skin of rabbits including a generic sample (accession number SAMN00013655), Rex black rabbit (accession number SAMN02693835), Rex white rabbit (accession number SAMN02693836) and Rex chinchilla rabbit (accession number SAMN02693834). Only 5 annotated genes (2 and 3 in Ocu1 and Ocu14

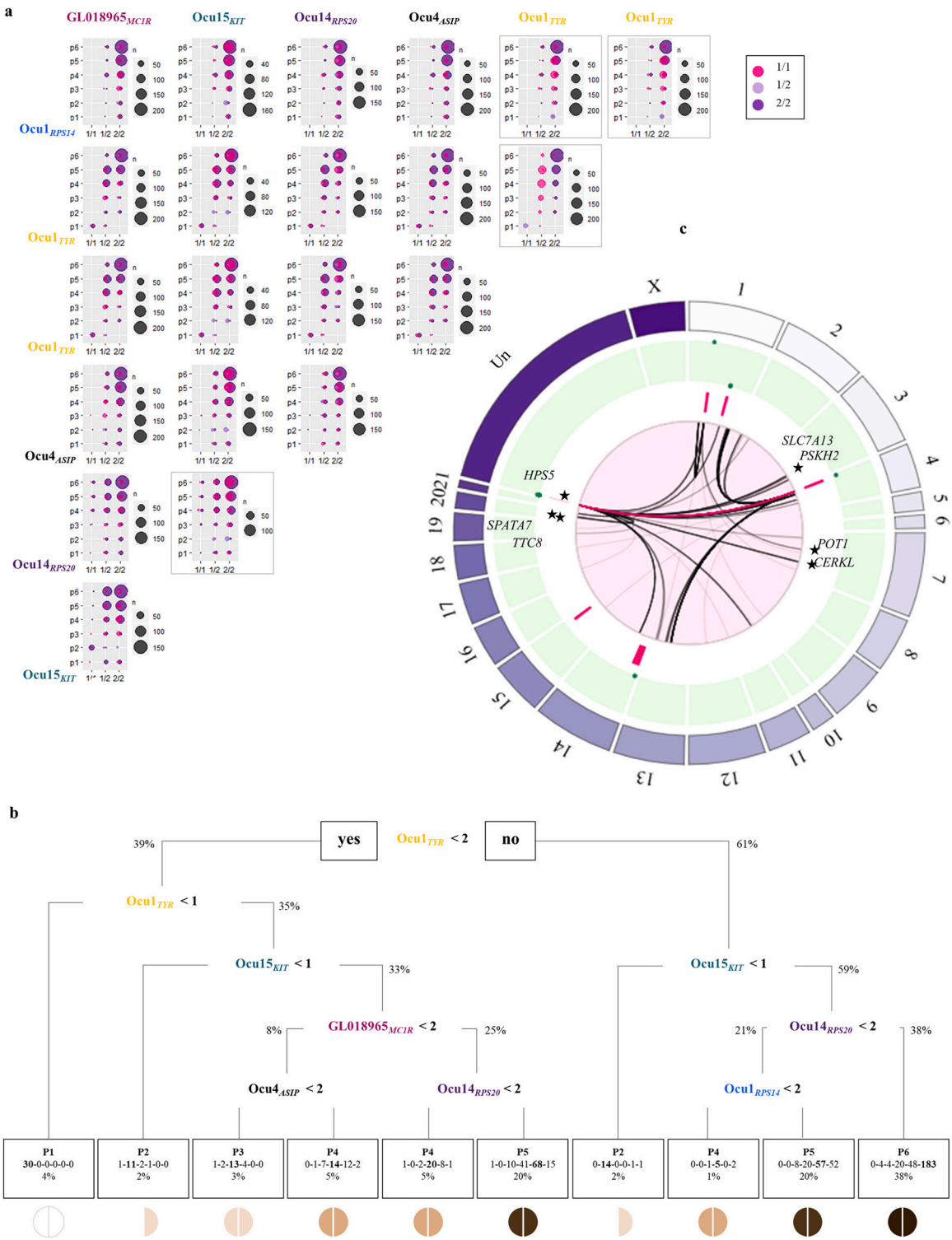

Fig. 3. Epistatic interactions underlying the molecular architecture of coat coloration. a. Pairwise epistatic interaction. Each count plot represents the genotypic distribution between the most significant markers of two out of the 7 associated loci previously identified. The x-axis represents genotypes at the locus mentioned on the line while the colored circles (pink, light purple and dark purple) represent genotypes at the locus mentioned on the column. The y-axis represents the 6 coat coloration phenotypic groups. The framed boxes correspond to significant interaction using the linear regression model. b. Decision tree. All rabbits were considered for all highlighted loci. The nodes discriminate genotypes at each involved marker with 0, 1 and 2 genotypes corresponding to the presence of 0, 1 or 2 minor alleles, respectively. The final leaves are the distinct colored phenotypes with the repartition of rabbits within each group from 1 to 6. c. Epistatic network. Only rabbits with colored phenotypes P3 to P6 were considered. Circular plot shows from external to internal tracks: chromosomes track (purple track), posterior inclusion probability (PIP) track ranging from 0 to 1 (green track), 6 identified regions track ($Ocu1_{RPS14}$, $Ocu1_{TYR}$, $Ocu4_{ASIP}$, $Ocu14_{RPS20}$, $Ocu15_{KIT}$ and $GL018965_{MC1R}$) (pink track) and interactions track (central circle). Stars represent significant interactions with genes already known to contribute to coloration process. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

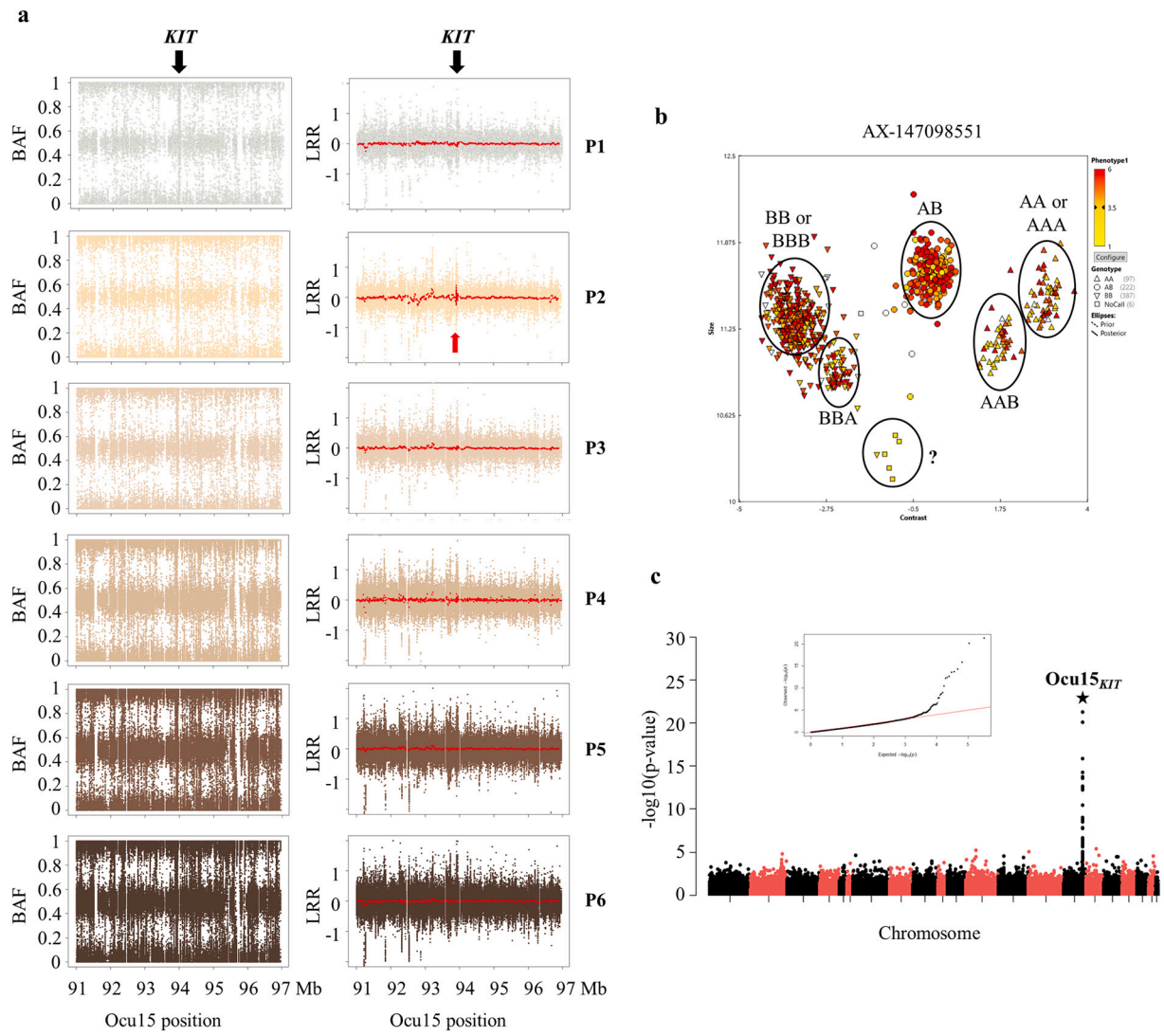

Fig. 4. Copy number variation within *KIT* associated with white spotting phenotype. a. B Allele Frequency (BAF) and Log R Ratio (LRR) plots. BAF and LRR values were plotted per phenotype (P1 to P6) for all individuals for a region of 6 Mb centered around the *KIT* gene. For each marker, the average LRR per phenotype was evaluated and a sliding average using a window of 10 markers was shown in red on LRR plots. b. Genotyping data extracted from the Axiom™ Analysis Suite Software 4.0.3.3 for the SNP AX-147098951 (93,916,801 bp) located within the *KIT* gene. Additional groups of genotypes likely reflecting a copy number variation spanning this marker. c. Manhattan plot. GWAS was performed using LRR values for each SNP and by a comparison of rabbits with phenotype P2 *vs.* all others groups of colored animals (P3 to P6). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

genomic regions, respectively) were quantified in skin including 3 ribosomal proteins pseudogenes (*RPS*) (Fig. 2b). It occurred that the three *RPS* pseudogenes (*RPS14* - ENSOCUG00000026323, *RPS27* - ENSOCUG00000024168 and *RPS20* - ENSOCUG00000025838) looked like processed pseudogenes since they all carried both START and STOP codons in the OryCun2.0 genomic reference sequence. While 8 and 10 pseudogenes of *RPS14* and *RPS20*, respectively are sparse in the rabbit genome, only the two copies located on Ocu1 and Ocu14, respectively are processed pseudogenes carrying the transcription initiation and ending codons (Fig. 2c). All of them were expressed in the skin tissue of the generic rabbit sample (Fig. 2b). Despite the lack of statistics between the three samples of skin of Rex rabbits (black *vs.* white *vs.* chinchilla), less reads mapped to both *RPS14* and *RPS20* processed pseudogenes in the black Rex rabbit (Fig. 2d).

3.2. Gene-gene interactions contribute to the determinism of coat color of body extremities

To assess whether specific interactions accounted for the variability of coat color of body extremities, we first evaluated pairwise genotypic distribution across the 6 phenotypic groups (P1 to P6) between the 7 selected markers from previous analyses (AX-146995791 ($Ocu1_{RPS14}$), AX-147087415 ($Ocu1_{TYR}$ [P1]), AX-147073566 ($Ocu1_{TYR}$ [P3 to P6]), AX-147097074 ($Ocu4_{ASIP}$), AX-147006836 ($Ocu14_{RPS20}$), AX-146983797 ($Ocu15_{KIT}$), and AX-147194100 ($GL018965_{MC1R}$). Significant interactions between variants located on Ocu1 were identified (Fig. 3a and Supplementary Table S6). Significant epistasis were also highlighted between best markers of $Ocu15_{KIT}$ and $Ocu14_{RPS20}$ (p-value = 0.0156) and a trend was observed for $Ocu4_{ASIP}$:$Ocu14_{RPS20}$ (p-value = 0.08127) and $Ocu4_{ASIP}$:$GL018965_{MC1R}$ (p-value = 0.1056) (Fig. 3a and Supplementary Table S6).

We built a classification tree based on the genotypes at each marker of this set of 7 markers to apprehend epistasis and genotypes-phenotypes relationships (Fig. 3b). As expected, $Ocu1_{TYR}$ and $Ocu15_{KIT}$ were found as major genes responsible of white (P1) and spotted (P2) phenotypes, respectively. For the remaining phenotypic groups, approximately 28%, 31%, 65% and 72% of individuals seemed correctly classified for phenotypes P3, P4, P5 and P6, respectively (Fig. 3b). While P3 and P4

phenotypes seem mostly explained by interaction between $Ocu4_{ASIP}$ and $GL018965_{MC1R}$, interactions between both loci involving *RPS* processed pseudogenes, $Ocu1_{RPS14}$ and $Ocu14_{RPS20}$, seemed involved in darker P5 and P6 phenotypes (Fig. 3b).

We evaluated the most significant genetic components including gene-gene interactions that contribute to the determinism of coat color of body extremities. We considered only phenotypes P3 to P6 but incorporated in our genetic model all possible combinations between the set of 7 selected variants. The best returned model included the different markers as main effect with the most significative positive effects for $Ocu4_{ASIP}$, $Ocu14_{RPS20}$ and $GL018965_{MC1R}$. Significant epistasis were highlighted with the most significant effect for the $Ocu4_{ASIP}$: $GL018965_{MC1R}$ interaction with a negative effect on phenotypes (Supplementary Fig. S6). In addition, $Ocu15_{KIT}$ showed significant epistatic effects with both $Ocu1_{TYR}$ and ribosomal genes *RPS* processed pseudogenes (Supplementary Fig. S6).

Finally, we considered epistasis between 6 of the 7 markers and the rest of the genome to analyse with a wide angle the coloration genes network. Only one variant, AX-147073566, was considered for $Ocu1_{TYR}$ since both selected markers (AX-147073566 and AX-147087415) are very close. The best interactions with a lfsr $<10^{-03}$ are shown on the Fig. 3c. Approximately 570 significant interactions were obtained with more than 85% of those involving $Ocu4_{ASIP}$ and $GL018965_{MC1R}$ markers with clusters of variants between each other (Supplementary Tables S7-S10). The best significant effect was observed for $Ocu4_{ASIP}$: $GL018965_{MC1R}$ interaction (lfsr $= 1.33 \times 10^{-10}$) with a similar effect to the previously one observed (Supplementary Table S9 and Supplementary Fig. S6). Interestingly, several novel epistatic interactions pinpointed to genomic regions spanning genes involved in coloration pathways or pigmentation linked disorders, such as *HPS5*, *POT1, TTC8, SPATA7* (Fig. 3c and Supplementary Tables S7-S10).

3.3. A copy number variation likely overlaps the $Ocu15_{KIT}$ locus

CNV affecting coat color genes were already associated with atypical coloration phenotypes in a broad range of species. We focused our research of CNV to regions spanning previously identified intervals as associated to phenotypes. Instead of the characterization of individual structural variants, we analysed the mean value of LRR (Log R Ratio) and BAF (B Allele Frequency) at each SNP of the region in the different phenotypic groups. For $Ocu15_{KIT}$, aberrant values of means of LRR, ranging below −0.97 and above 0.57, were observed for spotted rabbits (phenotype P2) in an interval containing the *KIT* gene, which referred to several markers that can then be considered as involved in a CNV (Fig. 4a). In addition, the distribution of BAF values for all individuals of the experimental design, not only the spotted colored rabbits, showed a typical profile of a CNV since we detected many values outside the expected 0, 0.5 and 1 categories (Fig. 4a). This suggests that the CNV affecting the *KIT* gene segregated within the whole protocol. The CNV seemed to affect the *KIT* gene as shown for 1 marker located within the gene for which the signal intensity seemed different between both alleles and additional groups of genotypes might be deduced (Fig. 4b). A similar pattern was observed for 5 markers located within the *KIT* gene suggesting a CNV of at least 5 Kb (93,911,613 bp - 93,916,801 bp) affecting the 5 last exons of the gene including the STOP codon. To confirm the strategy, animals that did not carry *a priori* the CNV based on BAF values 0 and 1 at the marker AX-147098551 and independently of their phenotypic group were considered as negative controls. The distribution of LRR values and the mean LRR did not show deviation from 0 in these individuals reinforcing the existence of a CNV at this location (Supplementary Fig. S7). A GWAS using individual LRR values as markers and comparing spotted colored animals (P2) to combined all others extremities-colored rabbits except P1 showed a significant signal on Ocu15 in the interval containing the *KIT* gene (Fig. 4c).

4. Discussion

4.1. Effect of known genes and/or mutations on the coat color of body extremities

The interval located on Ocu1 contains the *TYR* gene, which is an essential enzyme of the melanin biosynthesis from the tyrosine within melanosomes [9]. Regarding the Himalayan C^h allele, it perfectly discriminated white rabbits from all other individuals with coat color at their body extremities. Surprisingly, for the white phenotype, the best association signal under the recessive model was obtained for AX-147087415 instead of the Albino c allele. Four individuals classified in phenotype P1 are heterozygous for this known allele. Additional manual genotyping for this variant showed genotyping errors from the SNP array with 2 out of 4 animals homozygous for the Albino c allele. The remaining 2 incoherent rabbits are likely phenotypic errors. Interestingly, markers located within the genomic region spanning the *TYR* locus also showed association signals when only phenotypes P3 to P6 were considered. The LD structure measured in this region suggested the segregation of two distinct haplotypes. The haplotype carrying the region upstream the *TYR* gene may likely be involved in the variability of coat color of body extremities under an additive determinism. As observed on the classification tree, heterozygosity at the $Ocu1_{TYR}$ seemed more correlated with lighter P3 and P4 phenotypes while homozygosity for the major allele seemed more represented within darker P5 and P6 groups of rabbits. In a conditional and reversible gene expression knockdown mouse model, the authors showed that TYR was necessary not only for the synthesis of melanin, but also for the complete maturation of the stage IV melanosome [50]. This system likely suggested the potential effect of an additional genetic variation at the *TYR* locus [50].

The KIT receptor is also a key regulator activating the synthesis of eumelanin through the MAPK signalling pathway [9]. We figured out a structural variant spanning partially the *KIT* gene might contribute to the P2 phenotype. The *KIT* gene has been described in several species associated with coloration traits, especially with white spotting phenotype in cats [51], donkeys [52], camels [53], horses [54] and English spotting phenotype in rabbits [55]. In addition, structural variants involving *KIT* have been identified and associated with white spotting phenotypes such as in horses in which a heterozygous 1.9 Kb deletion represented a true null allele responsible of the depigmentation phenotype [56].

The region on Ocu4 spans the *ASIP* gene, a signalling ligand initiating the synthesis of pheomelanin pigment through its binding to the MC1R receptor [9]. In many species, the *ASIP* gene is involved in coloration traits and the best known is the agouti phenotype [57]. In rabbits, the causal mutation disrupting the protein discriminates between full and a dual coloration due to the expression of pheomelanin [27]. Here, results focusing on the *ASIP* locus are less clear since the best association signal was located more than 1 Mb downstream the agouti *a* marker with an intermediate level of LD suggesting another variant is involved in the light coat color of body extremities (phenotypes P3 and P4). More and more studies have highlighted structural variants encompassing or close to the *ASIP* gene and associated with coat colored phenotypes in different domestic species [58,59]. Indeed, a 11 Kb deletion affecting the *ASIP* gene was the most likely variant for the black and tan phenotype in rabbits [60].

The genomic region carrying the *MC1R* gene is also associated with coat color of body extremities in our experimental design. Although black and japanese alleles were genotyped [61,62], only the black dominant E^d allele segregated within our experimental design and was the best associated marker with coat color traits. Only 2 haplotypes spanning the whole GL018965 scaffold have been identified throughout the experimental design clearly suggesting either the causal role of the E^d allele or an additional mutation within the same haplotype.

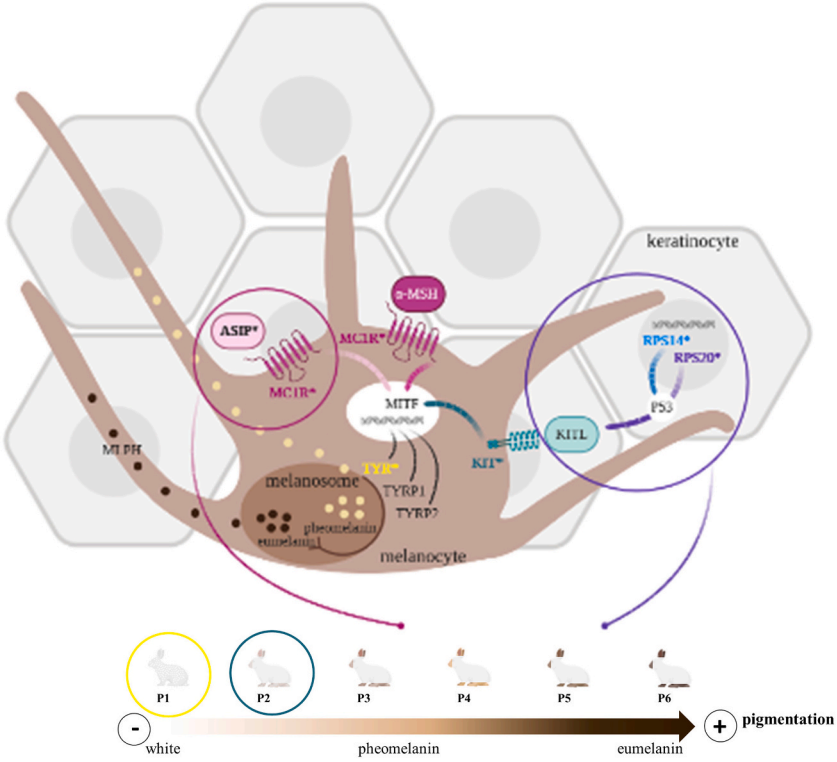

Fig. 5. Model of pigmentation molecular architecture through melanogenesis pathway. Both core genes and gene-gene interactions are represented. Core genes identified here to account for coat color variability are in bold, marked with a star and colored in yellow (*TYR*), green (*KIT*), dark pink (*MC1R*), light pink (*ASIP*), purple (*RPS20*) and blue (*RPS14*). Major genes involving *TYR* and *KIT* are responsible of white (P1) (yellow circle) and spotted (P2) (green circle) phenotypes, respectively. The remaining cream to dark brown phenotypes (P3 to P6) are explained by *ASIP*, *MC1R*, *TYR* and *RPS* genes as main effects but also through epistatic interactions. While *ASIP*:*MC1R* interactions contribute mainly to light colored rabbits (P3 and P4) (pink circle), *KIT*:*RPS* (including both *RPS14* and *RPS20*) epistasis seem account for darker colorations through tanning response (purple circle). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2. Processed pseudogenes of ribosomal proteins are likely involved into the coat color of body extremities

More importantly, two regions, Ocu1$_{RPS14}$ and Ocu14$_{RPS20}$, also account for the molecular architecture of the coat color of body extremities. Both regions contain few annotated genes on the OryCun2.0 genome and only some of them are expressed in the skin. Three of them belong to the 40S ribosomal proteins, which are likely the *RPS14*, *RPS27* and *RPS20* genes, respectively. The three ribosomal genes located within intervals of interest looked like pseudogenes which is very common for ribosomal genes in several species [63,64]. Analyses performed from publicly available RNA-seq data obtained from skin including a generic sample, a Rex black, a Rex white and a Rex chinchilla rabbit highlighted quantification of messengers from both *RPS14* and *RPS20* processed pseudogenes. The study of Tonner et al. detected transcription of ribosomal protein pseudogenes in diverse human tissues from RNA-seq data [66]. Unlike *RPS* genes that are constitutively expressed in almost all tissues, *RPS* pseudogenes are differentially expressed, suggesting that they may contribute to tissue-specific biological processes [66]. Two studies carried in mice [67] and zebrafish [68] showed coloration defects when mutations in *Rps20* and *Rps14*, respectively, were induced. Indeed, a study in mice reported mutations in *Rps19* and *Rps20,* leading to dark skin phenotypes, with a common signalling pathway through the stimulation of Kit ligand (kitl) expression by p53 [67]. Hence, a ribosome defect in keratinocytes may mimic ultraviolet response to keratinocytes resulting in a p53 induction in these cells that may drive melanocytes proliferation/migration *via* kitl signalling; this may lead to an hyperpigmentation tanning response [69,70]. In addition, deficiency in rps14 in zebrafish led to a delayed pigmentation through an increase of p53 activity [68]. Moreover, the comparison of the transcriptional profiles of human cell lines of dark and light melanocytes under basal conditions and following ultraviolet-B irradiation showed an interaction between ribosomal proteins and the p53 signalling pathway [71].

4.3. Epistatic network contributes to the genetic determinism of the coat color of body extremities

The most significant gene-gene interaction identified from the various analyses is between both markers tagging *ASIP* and *MC1R* loci. Epistasis seem to have a negative effect on coat coloration since combination of both variants is mainly associated with light colored rabbits harbouring cream/beige coat pigmentation (P3 and P4 phenotypes). Epistasis involving those genes has already been described in mice [72,73], human [14], sheep [74] but also in rabbits [27]. Interactions have been associated with color variation particularly in Creole sheep where it has clearly been showed that a functional wild-type genotype at *MC1R* locus is needed for the manifestation of the effects of the duplicated allele at the *ASIP* gene [74]. Although highlighting epistasis remains challenging, understanding the functional role of the interaction at the molecular level affecting phenotypes is still more complex. Consequence of *ASIP*:*MC1R* genetic interaction is straightforward since both proteins act within the same signalling pathway with ASIP being an antagonist ligand which competes with α-MSH for binding on its MC1R receptor. Our results also highlighted suggestive interactions between *RPS* processed pseudogenes themselves and *KIT*. Although the involvement of *RPS20* and *RPS14* processed pseudogenes needs to be assessed with further experiments, epistasis with those genes and its impact on pigmentation seemed consistent with knowledge as previously mentioned [67–69].

Beside epistasis highlighted between the hub genes involved in the determinism of coat coloration, we also pinpointed a denser interaction network including several genes, that are known to affect the pigmentation process. A significant interaction was identified between Ocu4$_{ASIP}$ locus and a part of the scaffold GL018733, in which the Heat Shock Protein 5 (*HPS5*) gene is located. Mutations within this gene alter melanosome biogenesis and have been associated with hypopigmentation specific of oculocutaneous albinism [75,76] and Hermansky-Pudlak Syndrome [76,77]. The built network also figured out *POT1* (The Protection Of Telomeres 1 protein) and *CERKL*, (Ceramide Kinase Like) located on Ocu7, significantly interacting with

GL018965$_{MC1R}$ locus. Several human genetics studies carried out in different ethnic groups have characterised mutations responsible for skin melanoma [78–80] making POT1 a major driver of this human disease as reviewed in [81,82].

5. Conclusions

The global overview of the coat color of body extremities determinism in our rabbit model (Fig. 5) showed the existence of an epistatic interaction network involving core genes but also likely additional genes with small effects. Our results suggested that *(i)* Ocu1$_{TYR}$ seemed to dictate whether pigmentation is produced or not and when the coloration occurred, it is restricted to body extremities, *(ii)* interactions among the others genes (Ocu1$_{RPS14}$, Ocu4$_{ASIP}$, Ocu14$_{RPS20}$ and GL018965$_{MC1R}$) seemed to dictate the amount of pigment produced and *(iii)* Ocu15$_{KIT}$ seemed to control where pigment is deposited, all body extremities or restricted to some extremities. The characterization of a genome-wide epistatic network might significantly contribute to a better understanding of underlying mechanisms. Moreover, divergences in the relationships between phenotypes and genotypes have been described in different breeds pointing out the functional effect of specific combination of alleles. Future studies through deeper analyses from sequencing data might lead to an allele-specific network considering also their dominance/recessivity or copy numbers.

Funding

This study has been funded by the European Union's H2020 project Feed-a-Gene under grant agreement no. 633531, the Animal Genetic division from INRAE, and the GenPhySE laboratory.

Authors' contributions

JD analysed data, performed visualisation and wrote the manuscript, YL performed quality control and analyses, NI organized sampling and genotyping, AD performed additional genotyping, SL performed phenotypic measurements and additional genotyping, HG contribute to statistical analyses, PA and FB were in charge of animal care and experimental design, JR set up the project, performed phenotypic measurements and quality control of datasets.

Availability of data and materials

The dataset are available in the repository (http://genoweb.toulouse.inra.fr/~jdemars/RabbitColoration/) belonging to the Genotoul Bioinformatics facility (http://bioinfo.genotoul.fr/).

Acknowledgements

We thank people of the animal facility, who carefully looked after the animals and for their help with the skin biopsies. We warmly thank L. Drouilhet, B. Servin and A. Vignal and all those who have contributed directly or indirectly to this work for discussion and support. We also thank both Centro Nacional de Genotipado (CeGen) (http://www.usc.es/cegen/) and GenoToul bioinformatics (http://bioinfo.genotoul.fr/) facilities.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2022.110361.

References

[1] M.E. Protas, N.H. Patel, Evolution of coloration patterns, Annu. Rev. Cell Dev. Biol. 24 (2008), https://doi.org/10.1146/annurev.cellbio.24.110707.175302.
[2] M. Cieslak, M. Reissmann, M. Hofreiter, A. Ludwig, Colours of domestication, Biol. Rev. 86 (2011) 4, https://doi.org/10.1111/j.1469-185X.2011.00177.x.
[3] I.C. Cuthill, W.L. Allen, K. Arbuckle, et al., The biology of color, Science (80-) 357 (6350) (2017), https://doi.org/10.1126/science.aan0221.
[4] J.T. Bagnara, J.D. Taylor, M.E. Hadley, The dermal chromatophore unit, J. Cell Biol. 38 (1) (1968), https://doi.org/10.1083/jcb.38.1.67.
[5] D.C. Bennett, M.L. Lamoreux, The color loci of mice - a genetic century, Pigment Cell Res. 16 (4) (2003), https://doi.org/10.1034/j.1600-0749.2003.00067.x.
[6] H.E. Hoekstra, Genetics, development and evolution of adaptive pigmentation in vertebrates, Heredity (Edinb). 97 (3) (2006), https://doi.org/10.1038/sj.hdy.6800861.
[7] L.M. San-Jose, A. Roulin, Genomics of coloration in natural animal populations, Philos Trans R Soc B Biol Sci. (2017), https://doi.org/10.1098/rstb.2016.0337.
[8] W.J. Pavan, R.A. Sturm, The genetics of human skin and hair pigmentation, Annu. Rev. Genomics Hum. Genet. (2019), https://doi.org/10.1146/annurev-genom-083118-015230.
[9] J.Y. Lin, D.E. Fisher, Melanocyte biology and skin pigmentation, Nature. (2007), https://doi.org/10.1038/nature05660.
[10] C.B. Kaelin, X. Xu, L.Z. Hong, et al., Specifying and sustaining pigmentation patterns in domestic and wild cats, Science (80-). 337 (6101) (2012), https://doi.org/10.1126/science.1220893.
[11] J. Pino, L. Kos, MC1R, EDNRB and kit signaling in pigmentation regulation related disorders, in: *Skin Pigmentation: Genetics, Geographic Variation and Disorders.*, 2013.
[12] F. Liu, M. Visser, D.L. Duffy, et al., Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up, Hum. Genet. 134 (8) (2015) 823–835, https://doi.org/10.1007/s00439-015-1559-0.
[13] S. Walsh, L. Chaitanya, K. Breslin, et al., Global skin colour prediction from DNA, Hum. Genet. 136 (7) (2017) 847–863, https://doi.org/10.1007/s00439-017-1808-5.
[14] M.D. Morgan, E. Pairo-Castineira, K. Rawlik, et al., Genome-wide study of hair colour in UK biobank explains most of the SNP heritability, Nat. Commun. 9 (1) (2018) 5271, https://doi.org/10.1038/s41467-018-07691-z.
[15] M.A. Carbone, A. Llopart, M. DeAngelis, J.A. Coyne, T.F.C. Mackay, Quantitative trait loci affecting the difference in pigmentation between Drosophila yakuba and D. santomea, Genetics. 171 (1) (2005), https://doi.org/10.1534/genetics.105.044412.
[16] N.I. Mundy, Coloration and the genetics of adaptation, PLoS Biol. 5 (9) (2007), https://doi.org/10.1371/journal.pbio.0050250.
[17] R.C. Albertson, K.E. Powder, Y. Hu, K.P. Coyle, R.B. Roberts, K.J. Parsons, Genetic basis of continuous variation in the levels and modular inheritance of pigmentation in cichlid fishes, Mol. Ecol. 23 (21) (2014), https://doi.org/10.1111/mec.12900.
[18] A. Wollstein, S. Walsh, F. Liu, et al., Novel quantitative pigmentation phenotyping enhances genetic association, epistasis, and prediction of human eye colour, Sci. Rep. 7 (2017), https://doi.org/10.1038/srep43359.
[19] M.D. Ritchie, K. Van Steen, The search for gene-gene interactions in genome-wide association studies: challenges in abundance of methods, practical considerations, and biological interpretation, Ann Transl Med. 6 (8) (2018) 157, https://doi.org/10.21037/atm.2018.04.05.
[20] W. Branicki, U. Brudnik, A. Wojas-Pelc, Interactions between HERC2, OCA2 and MC1R may influence human pigmentation phenotype, Ann. Hum. Genet. 73 (2) (2009) 160–170, https://doi.org/10.1111/j.1469-1809.2009.00504.x.
[21] S.G. Blanchard, C.O. Harris, O.R.R. Ittoop, et al., Agouti antagonism of Melanocortin binding and action in the B16F10 murine melanoma cell line, Biochemistry. (1995), https://doi.org/10.1021/bi00033a012.
[22] I. Suzuki, A. Tada, M.M. Ollmann, et al., Agouti signaling protein inhibits melanogenesis and the response of human melanocytes to α-melanotropin, J Invest Dermatol. (1997), https://doi.org/10.1111/1523-1747.ep12292572.
[24] L. Fontanesi, E. Scotti, D. Allain, Dall'Olio S., A frameshift mutation in the melanophilin gene causes the dilute coat colour in rabbit (Oryctolagus cuniculus) breeds, Anim. Genet. (2014), https://doi.org/10.1111/age.12104.
[25] V.J. Utzeri, A. Ribani, L. Fontanesi, A premature stop codon in the TYRP1 gene is associated with brown coat colour in the European rabbit (Oryctolagus cuniculus), Anim. Genet. (2014), https://doi.org/10.1111/age.12171.
[26] B. Aigner, U. Besenfelder, M. Müller, G. Brem, Tyrosinase gene variants in different rabbit strains, Mamm. Genome 11 (8) (2000) 700–702, https://doi.org/10.1007/s003350010120.
[27] L. Fontanesi, L. Forestier, D. Allain, et al., Characterization of the rabbit agouti signaling protein (ASIP) gene: transcripts and phylogenetic analyses and identification of the causative mutation of the nonagouti black coat colour, Genomics. (2010), https://doi.org/10.1016/j.ygeno.2009.11.003.
[28] C. Larzul, H. De Rochambeau, Selection for residual feed consumption in the rabbit, Livest. Prod. Sci. (2005), https://doi.org/10.1016/j.livprodsci.2004.12.007.
[29] L. Drouilhet, H. Gilbert, E. Balmisse, et al., Genetic parameters for two selection criteria for feed efficiency in rabbits, J. Anim. Sci. (2013), https://doi.org/10.2527/jas.2012-6176.
[30] H. Garreau, J. Ruesche, H. Gilbert, et al., Estimating direct genetic and maternal effects affecting rabbit growth and feed efficiency with a factorial design, J. Anim. Breed. Genet. (2019), https://doi.org/10.1111/jbg.12380.
[31] H. Wickham, Ggplot2 Elegant Graphics for Data Analysis (Use R!), 2016, https://doi.org/10.1007/978-0-387-98141-3.
[32] M. Carneiro, C.J. Rubin, F. Di Palma, et al., Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication, Science 345 (6200) (2014) 1074–1079, https://doi.org/10.1126/science.1253714.

[33] M. Sargolzaei, J.P. Chesnais, F.S. Schenkel, A new approach for efficient genotype imputation using information from relatives, BMC Genomics (2014), https://doi.org/10.1186/1471-2164-15-478.
[34] N.K. Kadri, G. Sahana, C. Charlier, et al., A 660-kb deletion with antagonistic effects on fertility and Milk production segregates at high frequency in Nordic red cattle: additional evidence for the common occurrence of balancing selection in livestock, PLoS Genet. 10 (1) (2014), e1004049, https://doi.org/10.1371/JOURNAL.PGEN.1004049.
[35] W. Drobik-Czwarno, A. Wolc, J.E. Fulton, J.C.M. Dekkers, Detection of copy number variations in brown and white layers based on genotyping panels with different densities 06 biological sciences 0604 genetics, Genet. Sel. Evol. 50 (1) (2018) 1–13, https://doi.org/10.1186/S12711-018-0428-4/TABLES/7.
[36] H. Thorvaldsdóttir, J.T. Robinson, J.P. Mesirov, Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration, Brief. Bioinform. 14 (2) (2013), https://doi.org/10.1093/bib/bbs017.
[37] X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies, Nat. Genet. (2012), https://doi.org/10.1038/ng.2310.
[38] X. Zhou, P. Carbonetto, M. Stephens, Polygenic modeling with Bayesian sparse linear mixed models, PLoS Genet. (2013), https://doi.org/10.1371/journal.pgen.1003264.
[39] D. Turner, S. Qqman, An R package for visualizing GWAS results using Q-Q and Manhattan plots, J Open Source Softw. (2018), https://doi.org/10.21105/joss.00731.
[40] G. Wang, A. Sarkar, P. Carbonetto, M. Stephens, A simple new approach to variable selection in regression, with application to genetic fine mapping, J. R. Stat. Soc. Ser. B Stat Methodol. (2020), https://doi.org/10.1111/rssb.12388.
[41] C.C. Chang, C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, J.J. Lee, Second-generation PLINK: rising to the challenge of larger and richer datasets, Gigascience. 4 (1) (2015) 7, https://doi.org/10.1186/s13742-015-0047-8.
[43] W.N. Venables, B.D. Ripley, Modern Applied Statistics with S Fourth Edition By, 2002, https://doi.org/10.2307/2685660.
[44] T. Therneau, B. Atkinson, B. Ripley, Package 'Rpart, 2015.
[45] M. Stephens, False discovery rates: a new deal, Biostatistics (2017), https://doi.org/10.1093/biostatistics/kxw041.
[46] Y. Cui, X. Chen, H. Luo, et al., BioCircos.Js: an interactive Circos JavaScript library for biological data visualization on web applications, Bioinformatics. (2016), https://doi.org/10.1093/bioinformatics/btw041.
[47] J.-H. Shin, S. Blay, J. Graham, B. McNeney, LDheatmap : an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms, J. Stat. Softw. (2006), https://doi.org/10.18637/jss.v016.c03.
[48] P. Behrouzi, E.C. Wit, De novo construction of polyploid linkage maps using discrete graphical models, Bioinformatics. (2019), https://doi.org/10.1093/bioinformatics/bty777.
[49] E. Cuthill, J. McKee, Reducing the bandwidth of sparse symmetric matrices, in: Proceedings of the 1969 24th National Conference, ACM 1969, 1969, https://doi.org/10.1145/800195.805928.
[50] E.K. Paterson, T.J. Fielder, G.R. MacGregor, et al., Tyrosinase depletion prevents the maturation of melanosomes in the mouse hair follicle, PLoS One 10 (11) (2015), https://doi.org/10.1371/journal.pone.0143702.
[51] M.P. Cooper, N. Fretwell, S.J. Bailey, L.A. Lyons, White spotting in the domestic cat (Felis catus) maps near KIT on feline chromosome B1, Anim. Genet. 37 (2) (2006) 163–165, https://doi.org/10.1111/j.1365-2052.2005.01389.x.
[52] B. Haase, S. Rieder, T. Leeb, Two variants in the *KIT* gene as candidate causative mutations for a dominant white and a white spotting phenotype in the donkey, Anim. Genet. 46 (3) (2015) 321–324, https://doi.org/10.1111/age.12282.
[53] H. Holl, R. Isaza, Y. Mohamoud, et al., A frameshift mutation in KIT is associated with white spotting in the Arabian camel, Genes (Basel) 8 (3) (2017), https://doi.org/10.3390/genes8030102.
[54] R. Hauswirth, R. Jude, B. Haase, et al., Novel variants in the KIT and PAX3 genes in horses with white-spotted coat colour phenotypes, Anim. Genet. 44 (6) (2013) 763–765, https://doi.org/10.1111/age.12057.
[55] L. Fontanesi, M. Vargiolu, E. Scotti, et al., The kit gene is associated with the english spotting coat color locus and congenital megacolon in checkered giant rabbits (oryctolagus cuniculus), PLoS One 9 (4) (2014), https://doi.org/10.1371/journal.pone.0093750.
[56] N. Dürig, R. Jude, H. Holl, et al., Whole genome sequencing reveals a novel deletion variant in the KIT gene in horses with white spotted coat colour phenotypes, Anim. Genet. 48 (4) (2017) 483–485, https://doi.org/10.1111/age.12556.
[57] W.K. Silvers, W.K. Silvers, The Agouti and extension series of alleles, Umbrous, and sable, The Coat Colors of Mice. (1979), https://doi.org/10.1007/978-1-4612-6164-3_2.
[58] B.J. Norris, V.A. Whan, A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep, Genome Res. (2008), https://doi.org/10.1101/gr.072090.107.
[59] A. Robic, M. Morisson, S. Leroux, et al., Two new structural mutations in the 5′ region of the ASIP gene cause diluted feather color phenotypes in Japanese quail, Genet. Sel. Evol. (2019), https://doi.org/10.1186/s12711-019-0458-6.
[60] A. Letko, B. Ammann, V. Jagannathan, et al., A deletion spanning the promoter and first exon of the hair cycle-specific ASIP transcript isoform in black and tan rabbits, Anim. Genet. 51 (1) (2020), https://doi.org/10.1111/age.12881.
[61] L. Fontanesi, M. Tazzoli, F. Beretti, V. Russo, Mutations in the melanocortin 1 receptor (MC1R) gene are associated with coat colours in the domestic rabbit (Oryctolagus cuniculus), Anim. Genet. (2006), https://doi.org/10.1111/j.1365-2052.2006.01494.x.
[62] L. Fontanesi, E. Scotti, M. Colombo, et al., A composite six bp in-frame deletion in the melanocortin 1 receptor (MC1R) gene is associated with the Japanese brindling coat colour in rabbits (Oryctolagus cuniculus), BMC Genet. (2010), https://doi.org/10.1186/1471-2156-11-59.
[63] Z. Zhang, P. Harrison, M. Gerstein, Identification and analysis of over 2000 ribosomal protein pseudogenes in the human genome, Genome Res. 12 (10) (2002) 1466–1482, https://doi.org/10.1101/gr.331902.
[64] S. Balasubramanian, D. Zheng, Y.J. Liu, et al., Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes, Genome Biol. 10 (1) (2009) R2, https://doi.org/10.1186/gb-2009-10-1-r2.
[66] P. Tonner, V. Srinivasasainagendra, S. Zhang, D. Zhi, Detecting transcription of ribosomal protein pseudogenes in diverse human tissues from RNA-seq data, BMC Genomics 13 (1) (2012), https://doi.org/10.1186/1471-2164-13-412.
[67] K.A. McGowan, J.Z. Li, C.Y. Park, et al., Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects, Nat. Genet. 40 (8) (2008) 963–970, https://doi.org/10.1038/ng.188.
[68] J. Ear, J. Hsueh, M. Nguyen, et al., A Zebrafish model of 5q-syndrome using CRISPR/Cas9 targeting RPS14 reveals a p53-independent and p53-dependent mechanism of Erythroid failure, J Genet Genomics. 43 (5) (2016) 307–318, https://doi.org/10.1016/j.jgg.2016.03.007.
[69] G. Walker, N. Box, Ribosomal stress, p53 activation and the tanning response, Expert. Rev. Dermatol. 3 (6) (2008) 649–656, https://doi.org/10.1586/17469872.3.6.649.
[70] H. Arnheiter, K. Bharti, Ribosomes and p53 - a new KIT for skin darkening, Pigment Cell Melanoma Res. 21 (5) (2008) 501–502, https://doi.org/10.1111/j.1755-148X.2008.00488.x.
[71] S. López, I. Smith-Zubiaga, A.G. De Galdeano, et al., Comparison of the transcriptional profiles of melanocytes from dark and light skinned individuals under basal conditions and following ultraviolet-B irradiation, PLoS One 10 (8) (2015), https://doi.org/10.1371/journal.pone.0134911.
[72] L.D. Siracusa, The agouti gene: turned on to yellow, Trends Genet. (1994), https://doi.org/10.1016/0168-9525(94)90112-0.
[73] C.C. Steiner, J.N. Weber, H.E. Hoekstra, Erratum: Adaptive variation in beach mice produced by two interacting pigmentation genes, PloS Biology 5 (9) (2007), https://doi.org/10.1371/journal.pbio.0050219 (PLoS Biol. 2008;6(2):0418. doi: 10.1371/journal.pbio.0060036).
[74] D. Hepp, G.L. Gonçalves, G.R.P. Moreira, T.R.O. de Freitas, Epistatic interaction of the Melanocortin 1 receptor and Agouti signaling protein genes modulates wool color in the Brazilian creole sheep, J Hered. 107 (6) (2016) 544–552, https://doi.org/10.1093/jhered/esw037.
[75] K. Okamura, T. Suzuki, Current landscape of Oculocutaneous albinism in Japan, Pigment Cell Melanoma Res. (2020), https://doi.org/10.1111/pcmr.12927.
[76] M. Mériot, C. Hitte, M. Rimbault, C. Dufaure de Citres, V. Gache, M. Abitbol, Donskoy cats as a new model of oculocutaneous albinism with the identification of a splice-site variant in Hermansky–Pudlak Syndrome 5 gene, Pigment Cell Melanoma Res. 33 (6) (2020) 814–825, https://doi.org/10.1111/pcmr.12906.
[77] C.M.S. Daly, J. Willer, R. Gregg, J.M. Gross, Snow white, a model of Hermansky-Pudlak syzndrome type 5, Genetics. 195 (2) (2013) 481–494, https://doi.org/10.1534/genetics.113.154898.
[78] J. Shi, X.R. Yang, B. Ballew, et al., Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma, Nat. Genet. 46 (5) (2014) 482–486, https://doi.org/10.1038/ng.2941.
[79] C.D. Robles-Espinoza, M. Harland, A.J. Ramsay, et al., POT1 loss-of-function variants predispose to familial melanoma, Nat. Genet. 46 (5) (2014) 478–481, https://doi.org/10.1038/ng.2947.
[80] L. Trigueros-Motos, Mutations in POT1 predispose to familial cutaneous malignant melanoma, Clin. Genet. 86 (3) (2014) 217–218, https://doi.org/10.1111/cge.12416.
[81] L. Larue, Centenary theme section, SKIN MALIGNANCIES SIGNIFICANCE. (2020), https://doi.org/10.2340/00015555-3494.
[82] J.A. Newton-Bishop, D.T. Bishop, M. Harland, Melanoma genomics, Acta Derm Venereol. 100 (2020) 266–271, https://doi.org/10.2340/00015555-3493, 100-year theme Skin malignancies.
[83] Leo Breiman, Jerome Friedman, Richard Olshen, Charles Stone. Classification And Regression Trees, 1st, Routledge, New York, 1984. https://doi.org/10.1201/9781315139470. (Accessed 25 October 2017).