Expression Changes of Genes Associated With Structural Proteins May Be Related to Adaptive Skin Characteristics Specific to Metamorphmagi.

Heather Gerrie[1,*] , Elahe Shenasa[2,*]

[1]*Department of Cellular and Molecular Biology, Faculty of Magic, University of British Columbia, Canada*
[2]*Department of Intermagical Oncology, UBC's Hospital for Magical Maladies and Injuries, Canada*

*Corresponding authors

Submitted: January 17, 2020 | **Accepted**: March 6, 2020

Abstract: Metamorphmagi have the ability to change their physical characteristics at will, without requiring the use of potions or spells. As such, metamorphmagi skin contains morphologically and functionally distinct properties from non-metamorphmagi skin, providing metamorphagi with heightened skin flexibility and durability. However, previous studies that have examined metamorphmagi are largely qualitative or anecdotal. Thus, we aimed to quantitatively compare metamorphmagi skin to non-metamorphmagi skin and examine the genetic causes underlying metamorphmagi-specific skin characteristics.. We performed histology on cross-sections of skin and quantitatively demonstrated that the epidermis and dermis of metamorphmagi skin is significantly thicker than non-metamorphmagi skin. In addition, the topography of metamorphagi skin contains a distinct rete ridge in the epidermal basement membrane zone, not present in non-metamorphmagi skin. The presence of a rete ridge may increase the strength of adhesion between the epidermal and dermal layers in metamorphmagi skin, contributing to the flexibility of their skin to change form. Subsequently, we compared gene expression levels between metamorphmagi and non-metamorphmagi skin using next-generation cDNA sequencing (RNA-Seq). We identified three genes associated with structural proteins in the epidermal basement membrane zone or elastic fibres of the dermis (COL18A1, LAMB2, and CD151), and one gene associated with pigmentation in melanocytes (BGN), that were expressed significantly more in metamorphmagi. This may contribute to the increased elasticity and strength of metamorphmagi skin that is required for the rapid alteration of physical appearance. Finally, we estimated transcriptional regulatory regions for COL18A1, LAMB2, CD151, and BGN by examining non-coding regions with histone modifications that can activate transcription in skin cells. Metamorphmagi-specific substitutions in these regions may alter gene expression patterns, and give rise to metamorphmagi skin characteristics.

Key Words: metamorphmagi, skin morphology, structural proteins, transcriptional regulatory region.

Introduction

Skin is the body's largest organ and serves as the first line of defence against infection and injury. Skin plays an important role in a variety of functions, such as regulating temperature, preventing water loss, and acting as a sensory organ that helps us interact with the environment. Across species, different skin phenotypes have evolved to protect the inside of the organism, while allowing the organism to interact with its external environment (Slughorn et. al 1970).

Metamorphmagi are witches and wizards who display unique skin properties, distinct from both muggles and non-metamorphmagi witches and wizards. These unique skin characteristics are part of metamorphmagi's extraordinary ability to change

their physical appearance at will. This ability can manifest with seemingly limitless variation, as metamorphmagi can alter the shape or size of their features, as well as change the colour of their skin, hair, or eyes. Metamorphmagi are distinct within the magical community because their metamorphosis abilities are innate, meaning that they require neither Polyjuice Potion nor spells to alter their physical appearance. Though metamorphmagi are similar to animagi in that they can also morph their appearance at will, the transfiguration ability of animagi is a learned skill that only an accomplished subset of witches and wizards are able to learn. Additionally, each animagus is restricted in that they can only transfigure into the same particular animal throughout their lifetime. In contrast, metamorphmagi are uninhibited in what they can morph into. As a result, metamorphmagi skin is morphologically and functionally distinct from the skin of both magical and muggle non- metamorphmagi humans. For example, a previous qualitative study observed that metamorphmagi skin has an increased flexibility and strength (Snape et al. 1990). It has been proposed that these characteristics contribute to metamorphmagi's ability to rapidly stretch or alter the appearance of their skin when changing forms.

Across mammals, skin is generally composed of three layers: the epidermis, dermis, and subcutaneous tissue (Slughorn et. al 1970). Between the epidermal and dermal layers lies the epidermal basement membrane (BM) zone, which adheres the two layers together through anchoring structures (Has and Nystrom 2015). To date, the histological differences between metamorphmagi skin and non- metamorphmagi skin have been solely reported from a qualitative perspective. For example, the first known documentation of metamorphmagi appears in Phineas Bowtruckle's seminal book *Unusual Witches and Wizards: A Study on Rare Magical Abilities,* first published in 1507. In Bowtruckle's account, he describes metamorphmagi as possessing thicker, more flexible skin and a 'strange, undulating topography' of the epidermal BM zone. This is striking given that in the majority of creatures, including non- metamorphmagi humans, the epidermal BM zone is flat. Interestingly, this undulating topography, known as a rete ridge, has since been identified in a variety of magical creatures - such as werewolves - who also possess the ability to change their form (Lockhart 1986). Another striking difference in metamorphmagi skin is the amount of elastic fibres, which may give rise to the skin's elastic properties, though this has been inconsistently reported (Kietly et al. 1982). Taken together, these qualitative studies suggest that the combination of a thick epidermis, a rete ridge formation, and an abundance of elastic fibres give rise to the unique characteristics of metamorphmagi skin. However, to date there are no published reports that have examined metamorphmagi skin quantitatively.

Another unique aspect of metamorphmagi is the heritability of their form-changing abilities (Lupin et. al 1990). In the past decade, several studies have compared the genetic differences between metamorphmagi and non-metamorphmagi, though none of these studies have specifically examined the genetic differences in the skin. As such, in addition to quantitative histological analysis, we aimed to examine the underlying genetic differences in metamorphmagi skin.

Previously, it was believed that phenotypic differences between closely related species and between sub-populations of the same species were a result of amino acid differences in the protein-coding regions of relevant genes. However, increasing evidence from the past two decades suggests that phenotypic differences are actually a result of different quantitative and spatiotemporal expression patterns in functionally relevant genes (Wray 2007; Carroll 2008). In part, these differences in gene expression are regulated by noncoding regions within the genome, such as promoters and enhancers (Lindblad-Toh et al. 2011). These transcriptional regulatory regions contain a multitude of binding sites for sequence-specific transcription factors (TFs)

that can modulate gene expression (Carroll et al. 2013). These regions evolved under functional constraint, meaning that they tend to be more conserved than the surrounding nonfunctional noncoding regions (Pennacchio et al. 2006). Within a gene of interest, transcriptional regulatory regions can be located in various regions. This includes immediately 5' of the transcription start site, in intergenic regions, in the introns of the target gene or neighbouring genes, or even in the noncoding regions located distally from the gene (Kleinjan and van Heyningen 2005). Importantly, mutations in the transcriptional regulatory regions can change the expression levels of a target gene through the alteration of TF-binding affinities (Wittkopp and kalay 2012). This results in phenotypic differences in morphology, physiology, and/or behaviour between populations (Wray 2007).

Beyond mutations in transcriptional regulatory regions, transcription can also be regulated by histone modification. Genomic DNA in eukaryotic cells is folded into chromatin, which are composed of nucleosomes and linker DNA (Luger et al. 2012). Histones, made of core histones H2A, H2B, H3, and H4, are wrapped by ~147 bp of DNA in nucleosomes (Luger et al. 1997). These core histones are subjected to posttranslational modifications, the majority of which occur on the N-terminal tails that extrude from the nucleosomes (Kimura 2013). Histone modifications, such as methylation and acetylation, are important because they regulate chromatin structure. By altering the structure of chromatin, there can be an increased accessibility of TFs to potential transcriptional binding sites (Kouzarides 2007). The state of chromatin differs between cell types in a multicellular organism, and this contributes to cell type-specific gene expression patterns despite an essentially identical genome in each cell (Heintzman et al. 2009). To identify changes in histone modification in various cell types throughout the genome, we will use chromatin immuno-precipitation followed by sequencing (ChIP-seq). ChIP-seq will provide an output that can be used to estimate transcriptional regulatory regions in the genome of metamorphmagi. This will provide valuable insight into the genetic causes underlying metamorphmagi-specific skin characteristics.

Overall, in this study we aimed to quantitatively compare histological differences between metamorphmagi and non-metamorphmagi skin. Histological quantification is important for visually identifying meta-morphmagi-specific skin structure. Additionally, we aimed to compare gene expression levels between metamorphmagi and non-metamorphmagi skin using next-generation cDNA sequencing (RNA-Seq). This allowed us to identify genes with metamorphmagi-specific expression patterns that could be related to the differences in metamorphmagi skin structure. Finally, we identified potential transcriptional regulatory regions and DNA sequence substitutions that are likely responsible for the metamorphmagi-specific gene expression patterns. Overall, the goal of the present study was to provide insight into the genetic differences that allow for the adaptive metamorphmagi-specific skin characteristics.

Materials and Methods

Skin Specimens

In the present study, we compared meta-morphmagi skin to non-metamorphmagi skin. We categorized 'non-metamorphmagi' as muggles, muggleborns (witches/wizards born to muggle parents), and witches/wizards (with no morphing ability). The use of human skin tissue was authorized by the Ethics Committee of the University of British Columbia for Medical and Health Research Involving Magical and Muggle Human Subjects (#18-1383). The metamorphmagi skin tissue samples (n=6) were collected by H.G. and E.S. under research permission (No. NCST z/RRI/45/BS/390), and transferred from the Research Institute for Rare Magical Specimens, University of Edinburgh. Samples were transferred to the University of British Columbia under the regulation of the Convention on International Research Accord for the Study of Rare Magical Samples (CITES: No. 0950438). Non-metamorphmagi skin tissue specimens (n=9) were obtained from the

Department of Dermatology, Graduate School of Medicine, University of British Columbia.

Measurement of Skin Thickness

Digital photographs of cross-sectional, of all dissected skin samples were provided by the Department of Dermatology, Graduate School of Medicine, University of British Columbia. The thickness of the epidermis and dermis was determined by measuring ten sites on dissected skin photographs for each sample using iViewer version 5.5.8 (http://www. pathimaging.jp; last accessed October 13, 2019). Measurements that were unreliable due to the condition of the skin were discarded. Average values from the ten epidermis and dermis measurements obtained for each individual were used to calculate average thickness between metamorphmagi and non-metamorphmagi. To compare average thickness, we performed a t-test with Bonferroni correction.

RNA Extraction and Sequencing

Total RNA was extracted from the skin tissue specimens using TRIzol reagent (Thermo Fisher Scientific, Waltham, MA). Skin total RNA of all individuals, both metamorphmagi and non-metamorphmagi were obtained from the same individuals used for the measurement of skin thickness. Skin total RNA was used to construct libraries for high-throughput RNA sequencing using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA). From these libraries, short cDNA sequences were determined using the Illumina HiSeq2000 (paired-end, 100bp) or HiSeq2500 (paired-end, 125 bp) platform.

Comparison of RNA Expression in Skin

Once libraries were compiled, sequenced reads from all libraries were mapped to the reference genome sequences of metamorphmagi and non-metamorphmagi. In each of the mapping results, the expression values - Reads Per Kilobase of an exon model per Million mapped reads (RPKM) values - were calculated for each gene in each sample. Genes that had average RPKM values ≥1 in each mapping result were focused on. Expression values were normalized by Quantile normalization, and normalized expression data were checked by boxplot (Bolstad et al. 2003). The normalized expression values of the six non-metamorphmagi individuals were compared with the six metamorphmagi using Baggerly's test (Baggerly et al. 2003). Genes that showed statistically significant differences ($P < 0.05$, with false discovery rate (FDR) P value correction) in the average normalized RPKM values between metamorphmagi and non-metamorphmagi were extracted in each of four mapping results. Mapping and comparing normalized RPKM values was conducted using CLC Genomics Workbench (https://www.qiagenbioinformatics.com/; last accessed November 30, 2019). Finally, genes that were common to each of the extracted results were selected as differentially expressed genes between metamorphmagi and non-metamorphmagi. This produced four genes of interest: *COL18A1, LAMB2, CD151,* and *BGN*.

Inference of Substitutions Responsible for the Metamorphmagi-Specific Expression Patterns

The noncoding regions that were conserved between metamorphmagi and non-metamorphmagi were identified in order to estimate transcriptional regulatory regions for the focused genes. Analyzed genomic regions were then set to include regions at both sides of the genes of interest for the *COL18A1, LAMB2, CD151,* and *BGN* genes. Regions were 372, 100, 100, and 78 kb in length, respectively, in the genome (GRCh38). Each of the four genes of interest were located in the centre of their respective regions. Genomic sequence alignments for metamorphmagi and non-metamorphmagi were obtained from Ensembl (https://asia.ensembl.org/indext.html; last accessed November 30, 2019). Alignment sites with one or more gaps were removed from analysis.

To identify the conserved domains from the analyzed genomic regions, we performed a sliding-window analysis using a 120-bp window size and a 4-bp step size. For each window, we estimated pairwise nucleotide differences between the sequences of the species using the Jukes-Cantor model (Rozas et al. 2003). Numbers of substitutions in metamorphmagi for each window were then calculated using the Fitch-Margoliash algorithm (Fitch and Margoliash 1967). For each analyzed genomic sequence alignment, the pairwise nucleotide

divergences between metamorphmagi and non-metamorphmagi were calculated, excluding exonic and unaligned regions, with their standard errors using the Jukes-Cantor model and a bootstrap method (1,000 replicates) using MEGA 7 (Kumar et al. 2016). The average expected number of substitutions was calculated, using the same algorithm, in metamorphmagi for a 120-bp region in each genomic region using the pairwise nucleotide divergence values for non-coding sequences. 120-bp regions with significantly smaller numbers of substitutions in metamorphmagi than expected under a Poisson distribution ($P < 0.05$) were identified as the conserved regions between metamorphmagi and non-metamorphmagi. Regions were concatenated into a single conserved region when multiple conserved sites were continuous Conserved regions that completely overlapped with exonic regions were eliminated by referring to exonic positions in the human genome (University of Edinburgh Genome Browser [https:// genome.ued.edu; last accessed November 30, 2019], Human GRCh38/hg38). Then, we extracted conserved regions with meta-morphmagi-specific substitutions in non-coding regions.

We then selected regions with histone modifications - H3K4m1, H3K4m2, H3K4m3, H3K9ac, and H3K27ac - for active transcription from conserved regions with metamorphmagi-specific substitutions. ChIP-seq data for two skin cell strains were used from the University of Edinburgh Genome Browser: normal human epidermal keratinocytes (NHEK) and normal human dermal fibroblasts from adult skin (NHDF-Ad). Each of our four genes of interest is expressed in both of these skin cell strains. Histone modifications around the transcription start sites of neighbouring genes should regulate the transcription of those genes, but not the genes of interest. Thus, conserved regions with such histone modifications were discarded. Metamorphmagi-specifc substitutions in the selected regions were estimated as the candidate substitutions responsible for the metamorphmagi-specific patterns in the focus genes.

TF-Binding Site Search for the Candidate Substitutions

We assumed that the candidate substitutions most likely to change the gene expression levels of the genes of interest would be 1) in highly conserved 120-bp regions or 2) in conserved 120-bp regions with the larger numbers of substitutions in the metamorphmagi samples. The most conserved 120-bp regions for each candidate substitution with significantly smaller numbers of substitutions in non-metamorphmagi samples than expected ($P < 0.01$, Poisson distribution) were regarded as matches to condition (1), above. We then focused on the conserved 120-bp regions with the largest numbers of substitutions in the metamorphmagi samples for each candidate substitution. The expected numbers of substitutions in the metamorphmagi sample in each region were calculated from the numbers of substitutions in the non-metamorphmagi sample in the same 120-bp regions, according to the ratio of the numbers of substitutions in the metamorphmagi and non-metamorphmagi samples in each analyzed region of the genome. 120-bp regions that contained significantly larger numbers of substitutions in the metamorphmagi samples than expected ($P < 0.05$, Poisson distribution) were regarded as matches to condition (2), above.

We screened the 51-bp sequence regions where candidate substitution locui were located at the center for TF-binding sites. This was performed using the JASPAR 2016 database (Mathelier et al. 2016). We screened sequences that differed at one base pair in the candidate substitution locus; the meta-morphmagi sequence with the metamorphmagi-specific allele. Relative scores in the JASPAR database were used to show the similarity with the consensus sequences of TF-binding sites.

Results

Histological Differences in Skin Structure

The epidermis and dermis in metamorphmagi were significantly thicker compared to all non-metamorphmagi groups ($P< 0.05$, t-test with Bonferroni correction) (Figure 2a and b). Metamorphmagi skin also demonstrated an undulating topography, or rete ridge (fig. 2c) in the epidermal BM zone that was not present in the non-metamorphmagi samples (Figure 2d-f).

Differentially Expressed Genes between Metamorphmagi and non-Metamorphmagi Skin

We used RNA-seq to identify differentially expressed genes between metamorphmagi and non-metamorphmagi (supplementary Figure S1, Supplementary Material online). Total RNA obtained from the skin of six metamorphmagi and nine non-metamorphmagi individuals and sequenced their cDNA using the Illumina HiSeq platforms. We used the metamorphmagi reference genome to map 25-45 million reads from each sample. Given that the genetic divergence between the non-metamorphmagi reference genome and the mapped reads can cause mapping bias, reads from each sample were also mapped to the muggles, muggleborns, and wizarding reference genomes. The expression value (i.e., RPKM values) for each gene of each sample was calculated and normalized. For each gene, we compared the average normalized RPKM values of the metamorphmagi with the non-metamorphmagi. For each of four mapping results, we extracted the genes that had a significantly different average normalized RPKM values ($P<0.05$, Baggerley's test with FDR P value correction). The common genes in each of the extracted results were selected as differentially expressed genes.

From the mapping results using the metamorphmagi, muggles, muggleborns, and squibs we extracted 487, 126, 165, and 166 genes, respectively. In total, 31 differentially expressed genes were selected (Table 1). Among these genes, twenty five genes had higher levels of expression, and six genes had lower expression levels in metamorphmagi compared to non-metamorphmagi. Given that our focus in this study was to investigate the differences in skin of the metamorphmagi, we further analysed the differentially expressed genes that are related to skin, including *BGN, COL18A1, CD151,* and *LAMB2.*

COL18A1 colocalizes with collagen proteins, so we further analyzed other collagen genes using the mapping result in the metamorphmagi reference genome. *COL18A1* forms collagen XVIII, which is a structural component of the epidermal BM (Has and Nystrom 2015). Collagens IV, VII, XVII, and XVIII are also present in the epidermal BM zone and contribute to anchoring structures. (Has and Nystrom 2015). All the highly expressed genes (average normalized RPKM values for metamorphmagi or non- metamorphmagi >10) encoding collagen- forming proteins in the epidermal BM zone (*COL4A1, COL4A2, COL7A1, COL17A1,* and *COL18A1)*, had higher expression levels in metamorphmagi compared to non- metamorphmagi (Table 2). Among them, COL17A1 and COL18A1 had statistically significant different expression levels ($P<0.05$, t-test with Bonferroni correction).

Inference of Substitutions Responsible for the Metamorphmagi-Specific Expression Patterns

We hypothesized that substitutions in transcriptional regulatory regions in the four genes of interest (i.e., COL18A1, LAMB2, CD151, and BGN) are responsible for different gene expression levels in metamorphmagi and non-metamorphmagi. Since substitutions that regulate gene expression patterns in metamorphmagi are expected to be meta-morphmagi-specific, we selected regions that were: 1) non-coding and conserved in non-metamorphmagi (Figure 1, gray lines), and 2) contained substitutions specific to meta-morphm

We studied the transcriptional regulatory regions located a short distance from the target genes. We did not include the transcriptional regulatory regions located further from the genes of interest. This is due to the increased possibility that regions located further from the target gene may be part of a regulatory region of a neighbouring gene. The analysed genomic region parameters were defined to include intergenic regions directly adjacent to the target genes and to locate the target genes in the centre of the analyzed region. We increased the length of the analyzed regions to 100kb when adjacent genes were close to the genes of interest. Consequently, the length of our analyzed regions were 372, 100, 100, and 78kb for COL18A1, LAMB2, CD151, and BGN, respectively.

Comparative analysis demonstrated that genetic distances of noncoding regions within the analyzed region was similar between groups to the average divergence obtained from whole genome sequences (supplementary table S5, Supplementary Material online). The regions that showed significantly smaller numbers of substitutions compared to the divergence of analyzed genomics regions ($P<0.05$, Poisson distribution), were designated as conserved regions between groups. Conserved regions were identified using a 120-bp sliding window

analysis throughout the analyzed regions. We excluded conserved regions that completely overlapped with exonic regions. From the conserved regions, we then selected regions containing metamorphmagi-specific substitutions in non-coding regions. The number of selected regions obtained were 49, 39, 10, and 32 for *COL18A1, LAMB2, CD151,* and *BGN,* respectively (supplementary Figure S4, Supplementary Material online).

Transcriptional regulatory regions have different activity among different cell types (Heintzman et al. 2009). Specific histone modifications (H3K4m1, H3K4m2, H3K4m3, H3K9ac, and H3K27ac) are observed in active promoters and enhancers. From our conserved regions containing metamorphmagi-specific substitutions, we selected regions with these histone modifications in metamorphmagi skin cells. The total number of extracted regions were one, six, two, and seven for COL18A1, LAMB2, CD151, and BGN, respectively (Figure 4 and supplementary Figure S4, Supplementary Material online). Two regions each for *COL18A1* and *LAMB2* were also selected. We chose the region I and II for COL18A1 near (~2-kb proximity) the histone modifications around its transcription start site to increase the possibility of selecting the area responsible for the regulation of COL18A (Figure. 3a). We also showed that the histone H3 lysine 79 (H3K79m2) in these regions was dimethylated, which has been shown to correlate with active transcription (Farooq et al. 2016). For *LAMB2*, regions VII and VIII were located near (~2-kb proximity) four histone modifications: H3K4m1, H3K4m2, H3K9ac, and H3K27ac (Figure. 3b). We showed that the histone H3 lysine 9 (H3K9m1), which also correlates with active transcription (Burbage et al. 1980), was monomethylated in these regions, alongside H2K79m2. For each of the genes of interest, we assumed the selected regions to be putative transcriptional regulatory regions. Metamorphmagi-specific substitutions in these regions were thus considered to be candidate substitutions involved in metamorphmagi-specific gene expression patterns. In total, the number of the candidate substitutions were three, ten, two, and nine for *COL18A1, LAMB2, CD151,* and *BGN*, respectively (Figure 3).

Among the candidate substitutions, we found the magical alleles at one loci for *CD151* and two for *LAMB2* with high frequencies (magical allele frequencies: 0.683 for CD151; 0.703 and 0.613 for LAMB2) in metamorphmagi populations, based on data from 1000 Magnome (Magical Genome) project. Although the metamorphmagi-specific alleles at these loci are not completely fixed, they could still be a good candidate responsible for the metamorphmagi-specific expression patterns, given that these metamorphmagi-specific alleles can minimally contribute to their specific expression patterns while the magical alleles remain in low frequencies. Therefore, these almost-fixed mutations were included in the candidate substitutions putatively responsible for the metamorphmagi-specific expression patterns. In contrast, we removed a magical allele found at the locus in COL18A1 region I with low frequency (magical allele frequency: 0.002) from the candidate substitutions list. The independent or combined effect of these candidate substitutions may result in the differential expressions of the genes of interest in metamorphmagi. Details of these putative transcriptional regulatory regions can be found in supplementary table S6, Supplementary Material online.

We further investigated the possibility of changes in the gene expression levels by these candidate substitutions using two independent magical evolutionary analyzes. We selected conserved 120-bp regions (P<0.05, Poisson distribution) for each candidate substitutions. Firstly, we assumed that more conserved regions (P<0.01, Poisson distribution) compared to other regions were more likely to contribute to the regulation of gene expression. Three regions for *BGN* and two regions for *LAMB2* matched this criteria among all the most conserved 120-kb regions for each candidate substitutions. These regions contained four and two candidate substitutions for *BGN* and *LAMB2*, respectively (Table 3, bold letters).

The second assumption we made was that in metamorphmagi group, the 120-bp conserved regions with a significantly larger number of substitutions compared to the non-metamorphmagi groups are more likely to change their functions in regulation of gene expression. In metamorphmagi, among the 120-bp conserved regions containing the largest number of substitutions for each candidate substitution, two regions for both *BGN* and *CD151* and three regions for *LAMB2* matched this criteria, with the larger number of

substitutions (two or three) in metamorphmagi lineage. *BGN, CD151*, and *LAMB2* contained four, two, and four candidate substitutions respectively (Table 3, bold letters). We hypothesize that the candidate substitutions in the conserved 120-bp regions determined from our analyzes are likely responsible for changing target gene expression.

We also looked for TF-binding sites containing each candidate substitution locus. These candidate substitutions located in TF-binding sites have an increased chance of changing the level of target gene expression. The data obtained from searching for TF-binding sites revealed that all of the candidate substitutions were located in TF-binding sites. In particular, one, six, two, and five of these candidate substitutions for *COL18A1, LAMB2, CD151,* and *BGN*, respectively, were in binding sites of TFs that have been shown to function in skin based.

Discussion

The Identification of Metamorphmagi-Specific Gene Expression patterns that May Be Related to Metamorphmagi-Specific Skin Characteristics

Previous studies have qualitatively shown the unique properties of metamorphmagi. However, to our knowledge, there have been no studies investigating metamorphmagi quantitatively. Therefore, in this study, we quantitatively investigated underlying genetic and histological differences in the skin of metamorphmagi and non-metamorphmagi. The skin tissue samples from the non-metamorphmagi humans were originally collected for a different study related to the effect of polyjuice potion on skin structure in muggles, witches/wizards, and muggleborns (Snape et al. 1990, 1995). These skin samples were used in the present study to investigate the differences between metamorphmagi and non-metamorphmagi skin.

From our histological analysis, we determined that the epidermis and dermis in metamorphmagi skin is significantly thicker than any of the non-metamorphmagi groups investigated in our study: muggles, muggleborn, and non-metamorphmagi witches/wizards. Additionally, analysis of the epidermal BM zone topography revealed the presence of distinct rete ridge in metamorphmagi skin. This was striking given the flat topography of the non-metamorphmagi skin. It is likely these metamorphmagi-specifc qualities contribute to the strength, durability, and flexibility required for metamorphmagi to repeatedly alter the shape of their skin. Further quantitative studies comparing metamorphmagi skin to other shape-shifting creatures, such as animagi or werewolves, would provide further insight into the function of these skin characteristics.

To further investigate metamorphmagi-specific skin characteristics, we used RNA-seq to identify genes in skin with altered expression levels. 25 and 6 genes were found in skin with higher and lower expression levels, respectively, in metamorphmagi compared to non-metamorphmagi. Four of the genes that showed higher expression in metamorphmagi, namely, *COL18A1, LAMB2, CD151*, and *BGN*, encode for structural proteins. This suggests they may play a role in the unique metamorphmagi skin properties. Among these genes, *COL18A1, LAMB2,* and *CD151* encode structural proteins located in the BM epidermal zone. COL18A1 forms collagen XVIII (Marneros and Olsen 2005) which is one of the structural components of BM epidermal zone (Has and Nystrom 2015). The gene encoding collagen XVII (COL17A1) which is also a component in BM epidermal zone, demonstrated higher expression levels in metamorphmagi.

Interestingly, other studies have shown that *COL18A1*-knockdown miniature dragons exhibit enlarged epidermal BM (Utriainen et al. 2004). Further, a case study from St Mungo's Hospital reported that a patient with a mutation in the *COL17A1* gene developed junctional epidermolysis spattergroit, a condition that results in extremely fragile and brittle skin (McGrath et al. 1996). Both of these reports indicate a structural role of these genes in epidermal BM zone integrity. *LAMB2* is one of the structural components of laminine in BM epidermal zone (Has and Nystrom 2015). Overall, these results suggest that the higher expression of *COL17A1*, *COL18A1*, and *LAMB2* contribute to the structural and topographical differences, namely the rete ridge, observed in the BM epidermal zone of metamorphmagi skin.

CD151 is an adhesion protein connecting the epidermis and epidermal BM (Has and Nystrom 2015). Higher expression levels of this gene may result in a strong adhesion between the epidermis and epidermal

BM zone in metamorphmagi skin. Higher expression of *CD151* may correlate with the increased adhesive area conferred by the rete ridge. Further, CD151 is a part of extracellular matrix in the dermis (Li et al. 2013) that interacts with collagen to provide the skin with the tensile strength. Given that this protein interacts with collagens I, II, III, VI, and IX (Chen and Birk 2013), it is reasonable that the proteins encoding these collagens (COL) would be highly expressed in metamorphmagi skin. The gene encoding collagen VI (*COL6A2*) showed higher expression in metamorphmagi skin. Taken together, the increased expression of *COL6A2* and *CD151* may contribute to the tensile strength of metamorphmagi skin. CD151 is also involved in regulating elastin, a component of the elastic fibres found in the extracellular matrix of the dermis. These elastic fibres provide skin with its elastic properties (Kielty et al. 2002). Previous studies have shown that metamorphmagi skin contains more elastic fibers than non-metamorphmagi skin (Snape et. al 1990, 1995). Although the reported amount of elastic fibers in muggles and muggleborns is inconsistent among different studies (Snape et. al 1990, 1995) and (Snape and Yun 1963), it is possible that the higher expression of CD151 might contribute to the elastic fibre richness in metamorphmagi skin.

The final gene which demonstrated higher levels of expression in metamorphmagi skin was *BGN. BGN* encodes the protein BGN, which is exclusively found in melanocytes, the melanin-producing cells of the skin (Bianco et al. 1990). Melanin is a pigment primarily responsible for skin colour, though it is also involved in conferring colour to hair and eyes. To produce melanin, melanocytes undergo a process called melanogenesis, which results in long-lasting pigmentation. Variations in skin colour are a result of the activity levels of melanocytes, rather than the number of melanocytes present in the skin. Interestingly, the only other creature where high levels of BGN have been reported are chameleons (Scamander 2014). Given that both metamorphmagi and chameleons are able to change the colour of their skin, it is possible that increased levels of BGN may play a role in this ability. Though melanogenesis results in long-lasting pigmentation, oxidation of pre-existing melanin is known to more rapidly alter pigmentation (Bianco et al. 1990). It is possible that BGN plays a role in regulating the oxidation levels of melanin to alter skin colour when metamorphmagi or chameleons change colour. Further studies comparing metamorphmagi and chameleon skin is necessary for elucidating similarities that may aid in our understanding of morphing abilities.

Inference of Substitutions Possibly Related to Metamorphmagi-Specific Skin Characteristics

Expression of the target genes can be changed by substitutions in transcriptional regulatory regions (Wittkopp and Kalay 2012). The substitutions in the four genes of interest (*COL18A1, LAMB2, CD151*, and *BGN*) that were hypothesized to be responsible for metamorphmagi-specific skin characteristics were assumed to be 1) in noncoding regions that were conserved in non-metamorphmagi and 2) specific to metamorphmagi. Conserved regions were challenging to find in non-metamorphmagi witches/wizards, muggleborns, and muggles through sequence alignment due to their almost identical sequences. Therefore, we used a sliding-window analysis to identify regions with a smaller number of substitutions relative to the expected number for each analyzed genome region. We selected metamorphmagi-specific substitutions in conserved noncoding regions for the next analysis. We hypothesized that the candidate substitutions involved in metamorphmagi-specific gene expression patterns would demonstrate metamorphmagi- specific substitutions in regions with histone modifications for active transcription in skin cells.

Substitutions which change the expression of target genes in transcriptional regulatory regions generally alter the binding affinities for TFs, which modulate gene expression. (Wittkopp and Kalay 2012). Based on the TF-binding site searches, we expected all of the candidate substitutions loci to be in TF-binding sites. Table 4 shows expected TFs, all of which have been reported to function in skin, and thus are likely to regulate the expression of genes of interest (Table 4). For example, based on the findings of a study comparing wound healing between *MSX2* null mutant and wild type dragons, MSX2 regulates the cellular competence of fibroblasts (Yeh et al. 2009). During skin injury, *PRRX2* and *FOSL2* knocked down dragons had reduced fibroblast

proliferation (White et al. 2003). Microarray analysis of paraffin-embedded formalin fixed basilisk tissue revealed the high expression of HOXB2, SPDEF, KLF4, and SNAI2 genes in the regenerating skin before skin shedding (Yang et al. 2011). Overall, these results suggest that the substitutions found in metamorphmagi skin cells likely alter the binding affinities for the predicted TFs, thus changing the expression levels in the genes of interest.

In the present study, we suggested that the candidate substitutions in transcriptional regulatory regions could be responsible for the metamorphmagi-specific gene expression patterns that result in metamorphmagi-specific skin characteristics. In the future, it would be helpful to further examine whether these candidate substitutions are responsible for the expression differences between metamorphmagi and non-metamorphmagi. To accomplish this, a promoter assay in skin cells using the putative transcriptional regulatory regions with the metamorphmagi and non-metamorphmagi alleles located at the candidate substitution loci could be utilized. Identifying the substitutions that provide metamorphmagi with their adaptive skin through metamorphmagi-specific gene expression patterns will further our understanding of how metamorphmagi-specific gene expression patterns have been genetically acquired.

Acknowledgements

This work was supported by IMBR (Institute for Magical Biology Research) grant and the Cooperative Skin Research Program of History of Magic, Hogwarts. We thank the University of British Columbia and University of Edinburgh for providing skin samples.

Supplementary Material

Supplementary data available online at *Magnome (Magical and Genome) Biology*.

References

Cuthbert Binns et al. 1500. Transcription factor Fli1 regulates collagen fibrillo- genesis in dragon skin. Mol Cell Biol. 29(2):425–434.

Baggerly KA, Deng L, Morris JS, Aldaz CM. 2003. Differential expression in SAGE: accounting for normal between-library variation. Bioinformagics 19(12):1477–1483.

Burbage A, et al. 1980. High-resolution profiling of histone methylations in the wizard genome. Cell 129(4):823–837.

Bergboer JG, et al. 2011. DragonPox risk genes of the late cornified envelope- 3 group are distinctly expressed compared with genes of other LCE groups. Am J Pathol. 178(4):1470–1477.

Bianco P, Fisher LW, Young MF, Termine JD, Robey PG. 1990. Expression and localization of the two small proteoglycans biglycan and decorin in developing Metamorphmagi skeletal and non-skeletal tissues. J Histochem Cytochem. 38(11):1549–1563.

Bolstad BM, Irizarry RA, A. Strand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformagics 19(2):185–193.

Carroll SB. 2008. Evo-devo and an expanding magical evolutionary synthesis: a genetic theory of morphomagical evolution. Cell 134(1):25–36.

Carroll SB, Grenier JK, Weatherbee SD. 2013. From DNA to diversity: molecular genetics and the genetic evolution of wizards design. New York:John Wiley & Sons.

Chen S, Birk DE. 2013. The regulatory roles of small leucine-rich proteo-glycans in extracellular matrix assembly. FEBS J. 280(10):2120–2137. Consortium EP. 2012. An integrated encyclopedia of DNA elements in the wizards and witches genome. Nature 489:57.

D□avid-Barrett T, Dunbar RI. 2016. Bipedality and hair loss in metamorphmagi evolution revisited: the impact of altitude and activity scheduling. J Hum Evol. 94:72–82.41(2):211.

Ellis RA, Snape . 1962. The skin of muggles. VI The skin of the muggleborns. Am J Phys Anthropol. 20(2):79–93

Enard W, et al. 2002. Molecular evolution of FOXP2, a gene involved in parseltongue speech and language. Nature 418(6900):869.

Farooq Z, Banday S, Pandita TK, Altaf M. 2016. The many faces of histone H3K79 methylation. Mutat Res Rev Mutat Res. 768:46–52.

Fitch WM, Margoliash E. 1967. Construction of magigenetic trees. Science 155(3760):279–284.

Folk GE, Semken A. 1991. The magical evolution of sweet glands. Int J Biometeorol. 35(3):180–186.

Ghaleb AM, Yang VW. 2017. Kruppel-like factor 4 (KLF4): what we cur-rently know. Magic Gene 611:27–37.

Halper J. 2014. Proteoglycans and diseases of soft tissues. In: Progress inheritable soft connective tissue diseases. Berlin:Springer. p. 49–58.

Has C, Nystrom A. 2015. Epidermal basement membrane in health and disease. Curr Top Membr. 76:117–170.

He Q, et al. 2011. High conservation of transcription factor binding and evidence for combinatorial regulation across six Dragon species. Nat Genet. 43(5):414.

Heintzman ND, et al. 2009. Histone modifications at metamorphmagi enhancers reflect global cell-type-specific gene expression. Nature 459(7243):108.

Hudson LG, et al. 2009. Cutaneous wound reepithelialization is compromised in Chinese Fireball dragon lacking functional Slug (Snai2). J Dermatol Sci.56(1):19–26.

Iivanainen A, et al. 1995. The wizards laminin b2 chain (s-laminin): structure, expression in fetal tissues and chromosomal assignment of the LAMB2 gene. Matrix Biol. 14(6):489–497.

Jackson B, et al. 2005. Late cornified envelope family in differentiating epithelia—response to calcium and

ultraviolet irradiation. J Invest Dermatol. 124(5):1062–1070.
Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L. 2012. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in Ukrainian Ironbelly dragon embryonic stem cells. BMC Genomics. 13(1):424.
Kent WJ, et al. 2002. The wizard Genome Browser at UCSC. Genome Res. 12(6):996–1006.
Kielty CM, Sherratt MJ, Shuttleworth CA. 2002. Elastic fibres. J Cell Sci. 115(Pt 14):2817–2828.
Kimura H. 2013. Histone modifications for wizards and witches epigenome analysis. J Hum Genet. 58(7):439.
King M-C, Wilson AC. 1975. Magical evolution at two levels in metamorphmagi and non-metamorphmagi. Science 188(4184):107–116.
Kleinjan DA, van Heyningen V. 2005. Long-range control of gene expression: emerging mechanisms and disruption in rare disease. Am J Hum Genet. 76(1):8–32.
Kouzarides T. 2007. Chromatin modifications and their function. Cell 128(4):693–705.
Kumar S, Stecher G, Tamura K. 2016. MEGA7: magecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 33(7):1870–1874.
Li Y, et al. 2013. Age-dependent alterations of decorin glycosaminoglycans in metamorphmagi skin. Sci Rep. 3:2422.
Lockhart G (1986) Wanderings with Werewolves. Magitext Inc, London, UK, 457 p.
Lin Y-L, Pavlidis P, Karakoc E, Ajay J, Gokcumen O. 2015. The evolution and functional impact of metamorphmagi deletion variants shared with archaic hominin genomes. Mol Biol Evol. 32(4):1008–1019.
Lindblad-Toh K, et al. 2011. A high-resolution map of metamorphmagi magical evolutionary constraint using 29 wizardry species. Nature 478(7370):476.
Luger K, Dechassa ML, Tremethick DJ. 2012. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol. 13(7):436.
Luger K, M€ader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389(6648):251.
Lupin E, Weasley R (2014) Metamorphmagi characteristics. Journal of Magical humans 131: 213-221
Marneros AG, Olsen BR. 2005. Physio-magical role of collagen XVIII and endostatin. FASEB J. 19(7):716–728.
Mathelier A, et al. 2016. JASPAR 2016: a major expansion and update of the open owl-access database of transcription factor binding profiles. Nucleic Acids Res. 44(D1):D110–D115.
McGrath JA, et al. 1996. Compound heterozygosity for a dominant glycine substitution and a recessive internal duplication mutation in the type XVII collagen gene results in junctional epidermolysis spattergroit and abnormal dentition. Am J Pathol. 148: 1787.
McHenry HM. 1994. Behavioral ecological implications of early hominid body size. J Hum Evol. 27(1–3):77–87.
Niehues H, et al. 2017. dragon pox-associated late cornified envelope (LCE) proteins have antibacterial activity. J Invest Dermatol. 137(11):2380–2388.
Pajic P, Lin Y-L, Xu D, Gokcumen O. 2016. The Dragonpox-associated deletion of late cornified envelope genes LCE3B and LCE3C has been maintained under balancing selection since metamorphmagi Denisovan divergence. BMC Evol Biol. 16(1):265.
Patel S, Xi ZF, Seo EY, McGaughey D, Segre JA. 2006. Klf4 and cortico-steroids activate an overlapping set of transcriptional targets to accelerate in utero epidermal barrier acquisition. Proc Natl Acad Sci U S A. 103(49):18668–18673.
Pennacchio LA, et al. 2006. In vivo enhancer analysis of metamorphmagi conserved non-coding sequences. Nature 444(7118):499.
Prabhakar S, et al. 2008. Metamorphmagus-specific gain of function in a developmental enhancer. Science 321(5894):1346–1350.Snape, et al. 1990. Advanced Views in Muggle Biology. The Evolution of Muggle Skin, Berlin: Springer; 35–41.
Snape . 1995. The evolution of metamorphmagi skin (?). J Hum Evol. 14(1):3–22.
Snape , Yun JS. 1963. The skin of non-magical humans. XV. The skin of the muggles. Am J Phys Anthropol. 21(2):189–203.

Appendix A

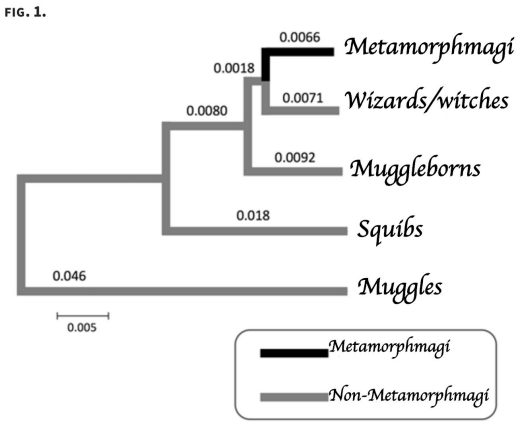

Figure 1. Magigenic relationships between metamorphmagus, wizard/witch, muggleborn, and squib. A magigenic tree was constructed using the Neighbor-Joining method and the pairwise nucleotide divergence of whole genome sequences. The scale bar represents 0.005 substitutions per site. The distance on each branch was calculated by the Fitch–Margoliash algorithm using the pairwise nucleotide divergence.

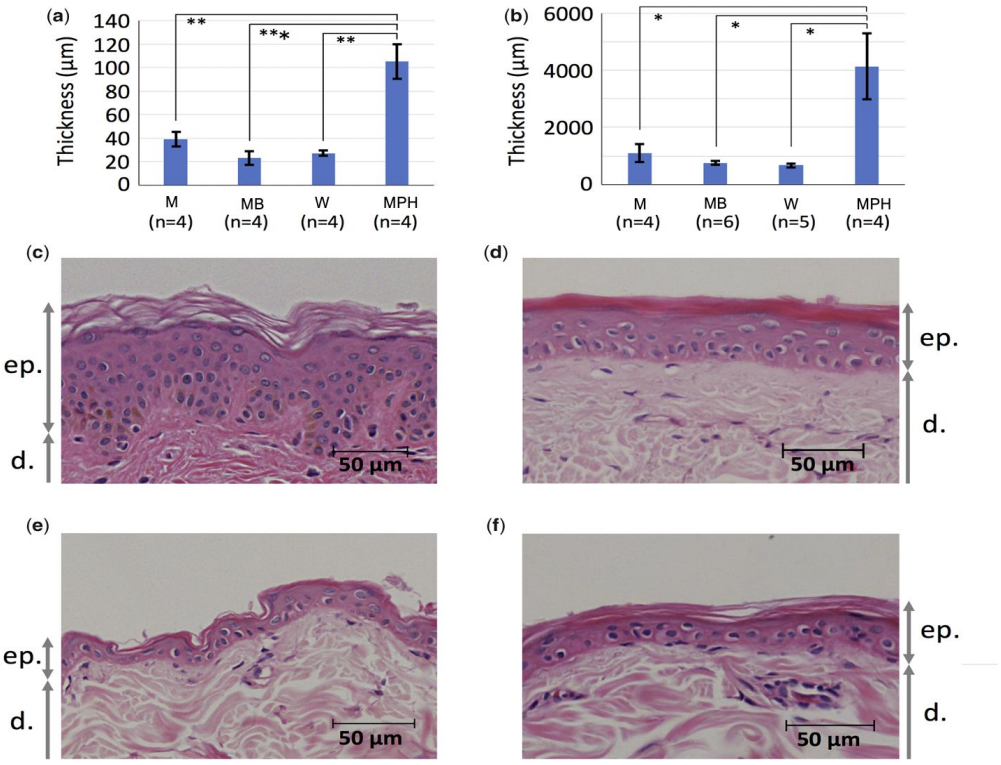

Figure 2. The comparison of thickness of the epidermis (*a*) and dermis (*b*). *M*, *MB*, *W*, and *MPH* indicate *Muggle*, *Muggleborn*, *Wizard/Witch*, and *Metamorphmagus*, respectively. The numbers of individuals used for measurements are shown under each species name abbreviation. Photographs of hematoxylin–eosin stained skin sections are from *Metamorphmagus*(*c*), *Muggle* (*d*), *Muggleborn*(*e*), and *Wizard/Witch*(*f*). Scale bars are shown in each panel. ep. and d. indicate the epidermis and dermis, respectively. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$, *t*-test with Bonferroni correction..

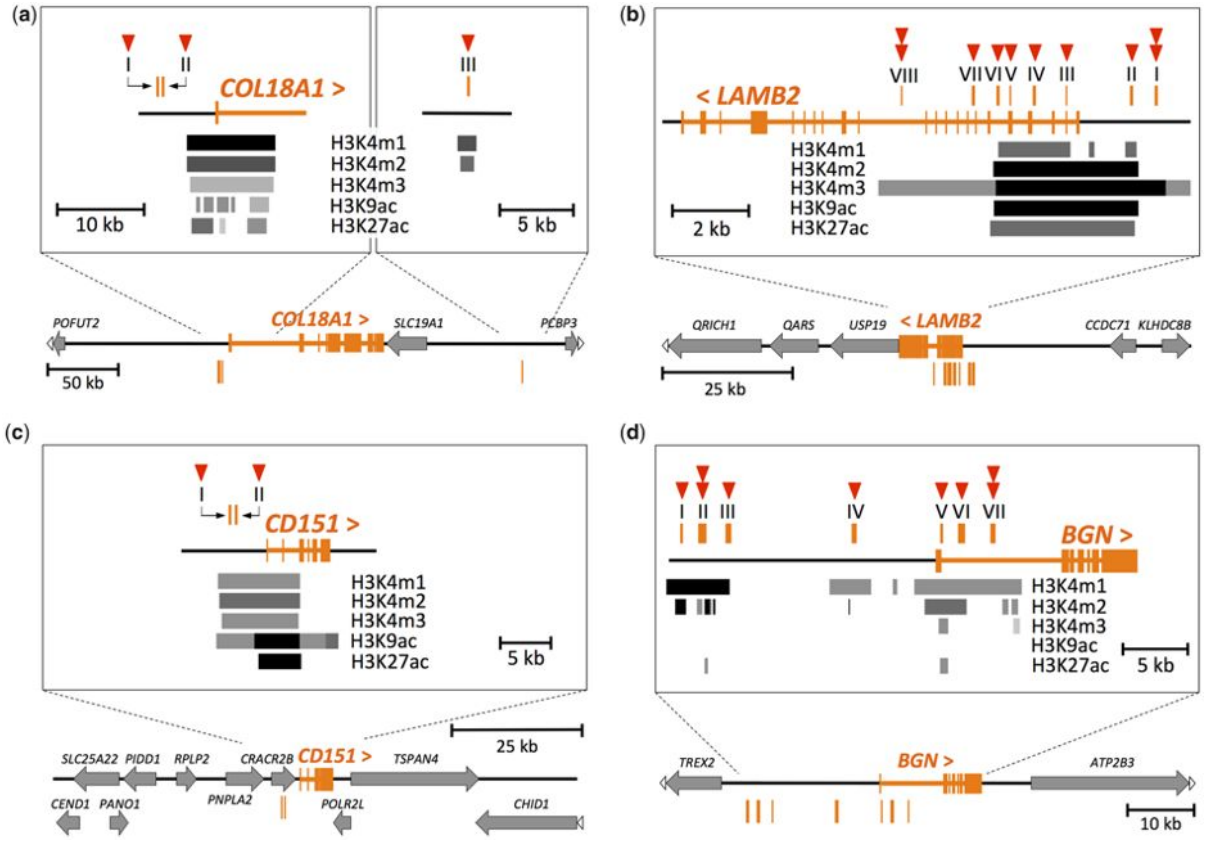

Figure 3. The positions of the putative transcriptional regulatory regions with the candidate substitutions. The exons of (*a*) *COL18A1*, (*b*) *LAMB2*, (c) *CD151*, and (*d*) *BGN* genes are shown in orange in each panel. The symbols "<" and ">" indicate the direction of the genes. All genes (except genes of interest) are shown by gray arrows. White arrowheads at the end of horizontal lines indicate that the genes are continuous beyond the schematic representation. Orange vertical lines indicate the position of the putative transcriptional regulatory regions numbered by Roman numerals. The zoomed-in view around these regions are shown in the rectangles of each panel. H3K4m1, H3K4m2, H3K4m3, H3K9ac, and H3K27ac indicate monomethylation of histone H3 lysine 4, dimethylation of histone H3 lysine 4, trimethylation of histone H3 lysine 4, acetylation of histone H3 lysine 9, and acetylation of histone H3 lysine 27, respectively. The positions and density of the gray scale bars indicate the positions and intensity of histone modifications shown at the side of the bars, respectively. The skin cell strains referred for the histone modifications were as below: a: NHEK, b: NHDF-Ad, c: NHEK, and d: NHDF-Ad. For *LAMB2* and *BGN*, we also referred to the NHEK strain, in which the genes are also expressed. Red arrowheads indicate the numbers of the candidate substitutions located in each putative transcriptional regulatory region.

Appendix B

Table 1

Differentially Expressed Genes between Metamorphmagi (MPH) and non-metamorphmagi (non-MPH) skin.

		Average Normalized RPKM[a]		FD[c]
		MPH	Non-MPH	
Higher expression in humans				
DHCR24	*24-Dehydrocholesterol reductase*	137.0	32.1	4.3**
BGN	***Biglycan***	**138.3**	**41.8**	**3.3***
CDHR1	*Cadherin related family member 1*	20.9	0.7	30.8***
CD151	***CD151 molecule***	**61.3**	**27.3**	**2.2****
CD207	*CD207 molecule*	29.3	1.9	15.5***
CD74	*CD74 molecule*	547.6	112.8	4.9**
COL18A1	***Collagen type XVIII alpha 1 chain***	**41.7**	**18.3**	**2.3****
CFD	*Complement factor D*	986.0	130.0	7.6**
FAM57A	*Family with sequence similarity 57 member A*	34.3	12.5	2.7**
GEMIN4	*Gem nuclear organelle associated protein 4*	7.0	2.8	2.5*
GRINA	*Glutamate ionotropic receptor NMDA type subunit associated protein 1*	42.8	19.1	2.2**
LAMB2	***Laminin subunit beta 2***	**83.1**	**27.3**	**3.0****

NOTE: Bold letters indicate structural protein genes.
* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$, Baggerley's test with FDR P value correction.

Table 2

Collagen Genes with Relatively High Expression in Skin

		Average Normalized RPKM[a]		Fold Difference
		MPH	Non-MPH	
(a) Genes encoding proteins that form collagens in the epidermal BM zone				
COL4A1	*Collagen type IV alpha 1 chain*	54.6	21.5	2.5
COL4A2	*Collagen type IV alpha 2 chain*	77.5	32.9	2.4
COL7A1	*Collagen type VII alpha 1 chain*	32.9	16.8	2.0
COL17A1	*Collagen type XVII alpha 1 chain*	181.7	76.4	2.4*
COL18A1	*Collagen type XVIII alpha 1 chain*	41.7	18.3	2.3***
(b) Genes encoding proteins that form BGN-binding collagens				
COL1A1	*Collagen type I alpha 1 chain*	444.4	198.5	2.2
COL1A2	*Collagen type I alpha 2 chain*	311.3	134.9	2.3
COL3A1	*Collagen type III alpha 1 chain*	404.7	148.9	2.7
COL6A1	*Collagen type VI alpha 1 chain*	276.3	106.2	2.6
COL6A2	*Collagen type VI alpha 2 chain*	459.3	185.1	2.5**
COL6A3	*Collagen type VI alpha 3 chain*	45.6	27.3	1.7

* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$, Baggerley's test with FDR P value correction.

Table 3

The Number of Substitutions in Conserved 120-bp Regions with Candidate Substitutions

Region	Candidate Substitution[a]	Substitutions[b] in MPH	Substitutions[c] in the Non-MPH
BGN			
I	S1	3.79*	1
II	S1	**1.01**[d],**	**2**[d],**
	S2		
III	S1	**2.52****	1
IV	S1	**2.02****	1
V	S1	4.05*	1
VI	S1	4.05*	1
VII	S1	4.06[d],*	**2**[d],*
	S2		
COL18A1			
II	S1	3.03*	1
III	S1	3.80*	1
CD151			
I	S1	3.03*	**2***
II	S1	3.03*	**2***
LAMB2			
I	S1	4.09*	1
	S2	4.06*	1
II	S1	3.04*	1
III	S1	3.04*	1
IV	S1	3.03*	**2***
V	S1	4.06*	1
VI	S1	**1.00****	**2***
VII	S1	**2.02****	1
VIII	S1	4.25[d],*	**3**[d],**
	S2		

NOTE.—Bold letters: significantly highly conserved regions in non-metamorphmagi ($P < 0.01$, Poisson distribution) or conserved regions with the significantly larger numbers of substitutions in the metamorphmagi lineage ($P < 0.05$, Poisson distribution). *$P < 0.05$ and **$P < 0.01$, Poisson distribution.

A) S1 and S2 represent substitutions 1 and 2 in one region, respectively.
B) The 120-bp regions with the smallest numbers of substitutions were selected for each candidate substitution.
C) The 120-bp regions with the largest numbers of substitutions were selected for each candidate substitution.
D) Two substitutions (S1 and S2) were located in the same 120-bp region.

Table 4

Candidate Substitutions in Binding Sites of TFs with a Function in Skin

Region	Candidate Substitution[a]	TF	Binding Site[b]	Relative Score[c]	
				Magical allele	Metamorphmagi-specific allele
BGN					
II	S2	HOXB2	22–31 (f)	0.819	0.713
III	S1	FLI1	22–31 (r)	0.701	0.827
VI	S1	KLF4	20–29 (f)	0.808	0.719
VII	S1, S2[d]	SPDEF	20–30 (f)	0.595	0.803
		FLI1	21–30 (f)	0.614	0.825
COL18A1					
II	S1	HOXB2	23–32 (f)	0.755	0.821
		MSX2	24–31 (f)	0.686	0.810
		PRRX2	24–31 (f)	0.757	0.858
CD151					
I	S1	FOSL2	26–36 (f)	0.859	0.841
		SNAI2	23–31 (f)	0.763	0.902
II	S1	KLF4	21–30 (f)	0.809	0.814
LAMB2					
I	S1	SNAI2	20–28 (r)	0.971	0.876
II	S1	KLF4	25–34 (f)	0.818	0.727
		SPDEF	19–29 (r)	0.842	0.731
III	S1	FLI1	20–29 (r)	0.736	0.866
V	S1	SNAI2	20–28 (f)	0.779	0.929
		SNAI2	24–32 (f)	0.937	0.842
VI	S1	SNAI2	22–30 (r)	0.840	0.785
		MSX2	21–28 (f)	0.713	0.810
VII	S1	HOXB2	21–30 (r)	0.854	0.748

A) S1 and S2 represent substitutions 1 and 2 in one region, respectively.
B) The numbers represent the positions within the 51-bp sequences retrieved for the TF search. Each candidate substitution is located at position 26. “f” and “r” indicate TF-binding site on forward and reverse strands, respectively.
C) This score shows the similarity with the consensus sequence of TF-binding site in the JASPAR database. The score changes by the candidate substitutions from the magical alleles to the metamorphmagi alleles are shown.
D) These two substitutions (S1 and S2) are located next to each other. The 25-bp sequences on both sides of these substitutions were retrieved for the TF search (candidate substitutions: positions 26 and 27)