Stem Cell Reports
Resource
OPEN ACCESS

Dynamic Transcriptional and Epigenetic Regulation of Human Epidermal Keratinocyte Differentiation

Alessia Cavazza,[1] Annarita Miccio,[2] Oriana Romano,[3] Luca Petiti,[4] Guidantonio Malagoli Tagliazucchi,[3] Clelia Peano,[4] Marco Severgnini,[4] Ermanno Rizzi,[4] Gianluca De Bellis,[4] Silvio Bicciato,[3] and Fulvio Mavilio[3,5,*]

[1]Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA
[2]Imagine Institute, 75015 Paris, France
[3]Department of Life Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy
[4]Institute for Biomedical Technologies, National Research Council, 20132 Milan, Italy
[5]Genethon, 1bis rue de l'Internationale, 91002 Evry, France
*Correspondence: fmavilio@genethon.fr
http://dx.doi.org/10.1016/j.stemcr.2016.03.003

SUMMARY

Human skin is maintained by the differentiation and maturation of interfollicular stem and progenitors cells. We used DeepCAGE, genome-wide profiling of histone modifications and retroviral integration analysis, to map transcripts, promoters, enhancers, and super-enhancers (SEs) in prospectively isolated keratinocytes and transit-amplifying progenitors, and retrospectively defined keratinocyte stem cells. We show that >95% of the active promoters are in common and differentially regulated in progenitors and differentiated keratinocytes, while approximately half of the enhancers and SEs are stage specific and account for most of the epigenetic changes occurring during differentiation. Transcription factor (TF) motif identification and correlation with TF binding site maps allowed the identification of TF circuitries acting on enhancers and SEs during differentiation. Overall, our study provides a broad, genome-wide description of chromatin dynamics and differential enhancer and promoter usage during epithelial differentiation, and describes a novel approach to identify active regulatory elements in rare stem cell populations.

INTRODUCTION

The epidermis is a stratified epithelium differentiating from keratinocyte stem cells (KSCs) contained in the basal layer and in the bulge of hair follicles. Upon division, KSCs produce transit-amplifying (TA) progenitors that generate differentiated keratinocytes and other epithelial skin components. The available information on the molecular events underlying self-renewing and differentiation of KSCs comes from studies on the murine hair follicle (reviewed in Blanpain et al., 2007). Much less is known about human KSCs, which lack robust markers for prospective isolation and are defined only retrospectively by the nature of their progeny in cell culture or transplantation assays. Clonal analysis in vitro has defined three types of clonogenic cells, giving rise to the so-called holoclones, meroclones, and paraclones. Holoclone-forming cells have the highest self-renewing and proliferative capacity, and define in culture the KSCs of the epidermis or the corneal epithelium (Pellegrini et al., 1999; Rochat et al., 1994). Meroclone- and paraclone-forming cells have proportionally less proliferative capacity and terminally differentiate into keratinocytes after 5–15 cell doublings, as expected for TA progenitors (Barrandon and Green, 1987). Few molecular markers are known for KSCs or TA progenitors: they include the p63, BMI1, CEBPs, MYC, and GATA-3 transcription factors (TFs), integrins, Wnt/β-catenin, NOTCH, HH, SGK3, and some bone morphogenetic proteins (Blanpain et al., 2007). In particular, p63 is considered a master regulator of morphogenesis, identity, and regenerative capacity of stratified epithelia (Pellegrini et al., 2001; Yang et al., 1999). Although some of the targets of p63 and other TFs involved in epidermal cell functions are known, little is known about the chromatin dynamics and the differential usage of promoters and enhancers driving the differentiation of human KSCs and TA progenitors.

Specific histone modifications are currently used to define chromatin regions with different regulatory functions. In particular, monomethylation of lysine 4 of histone 3 (H3K4me1) characterizes enhancer regions, whereas its trimethylation (H3K4me3) defines promoters (Ernst et al., 2011; Heintzman et al., 2009). Acetylation of H3K27 defines transcriptionally active enhancers and large clusters of enhancers (super-enhancers [SEs]) involved in the definition of cell and tissue identity (Hnisz et al., 2013). In this study, we aimed to map transcriptional regulatory elements and define their usage during epithelial differentiation. By combining high-throughput identification of Pol-II-transcribed (capped) RNAs defined by Cap Analysis of Gene Expression (DeepCAGE) (Carninci et al., 2006) with genome-wide profiling of histone modifications determined by chromatin immunoprecipitation (ChIP-seq), we mapped active enhancer and SE elements in prospectively isolated TA progenitors and terminally differentiated keratinocytes. For KSCs, which

lack markers for prospective isolation, we exploited the integration characteristics of the Moloney murine leukemia retrovirus (MLV), which integrates in active promoters and enhancers (Biasco et al., 2011; Cattoglio et al., 2010; De Ravin et al., 2014) as a consequence of the direct binding of the viral integrase to the bromodomain and extra-terminal (BET) proteins BRD2, BRD3, and BRD4 that tether the pre-integration complex to acetylated chromatin regions (De Rijck et al., 2013; Gupta et al., 2013; Sharma et al., 2013). By using MLV vector integration clusters as surrogate genetic markers of active regulatory elements, we mapped a collection of putative enhancers and SEs active in bona fide KSCs, retrospectively defined by their capacity to maintain long-term keratinocyte cultures.

RESULTS

DeepCAGE Mapping of Active Promoters in Keratinocyte Progenitors and Differentiated Keratinocytes

To enrich keratinocyte progenitors (KPs) from a keratinocyte mass culture, we panned β1 integrin-positive cells by adherence to collagen-IV-coated plates (Jones and Watt, 1993). Adhering cells were highly enriched in KPs, as determined by a clonogenic assay, and showed significantly increased expression of the progenitor-related markers *TP63* ($p < 0.05$), *LRIG1* ($p < 0.01$), *ITGB1*, *MCSP*, and *DLL1* ($p < 0.001$) by real-time qPCR, while the non-adhering fraction was depleted in colony-forming cells and expressed the differentiation markers *KRT1*, *IVL*, and *LOR* (Figures S1A–S1D). Differentiated keratinocytes (DKs) were obtained by in vitro differentiation in conditions of contact inhibition (Kouwenhoven et al., 2015; Shen et al., 2013), and showed residual colony-forming capacity and high expression of differentiation markers (Figures S1E and S1F).

To define global promoter usage, we used DeepCAGE on RNA extracted from KPs and re-analyzed an epidermal keratinocytes dataset available from ENCODE as a proxy of DKs. We identified a total of 15,283 CAGE promoters, 14,565 expressed in KPs and 15,027 in DKs. Most CAGE promoters mapped to known promoters (20%) or to immediately downstream 5′ UTR regions (48.6%) or gene bodies (Figure 1A). We grouped CAGE promoters in three clusters based on the tag position with respect to transcription start sites (TSSs): promoters in cluster 3 showed a broad profile around TSSs and represented the majority of alternatively used promoters, cluster 2 represented canonical promoters with a sharp localization at TSSs, while cluster 1 exemplified pervasive transcription within genes (Figure 1B).

Epithelial Differentiation Is Characterized by Quantitative Regulation of a Large Set of Common Promoters

Most CAGE promoters (14,309) were active in both cell populations and represented 98.2% and 95.2% of KP and DK promoters, respectively. Only 256 and 718 promoters were strictly stage specific, the majority of which (>60%) represented uncharacterized TSSs or were associated with non-coding transcripts, mainly long non-coding RNAs (lncRNAs). Most of the changes in transcriptome associated with keratinocyte differentiation were therefore defined by quantitative changes in the expression of promoters active in both KPs and DKs. A total of 5,429 promoters were expressed at significantly different levels between KPs and DKs ($p < 0.001$, χ^2 test), with 1,838 promoters upregulated in KPs and 1,712 in DKs at a $\log_2$ fold change (FC) of ≥ 2 (Figure 1C). In KPs differentially expressed TSSs were more abundant in promoters and introns, while in DKs they were more abundant in introns and 3′ UTRs (Figure 1D). qPCR analysis confirmed differential mRNA expression for 40 of the 46 randomly chosen promoters (Figure S2A). We detected alternative transcription initiation in 1,187 protein-coding genes, 455 of which underwent switch between alternative promoters during the KP-to-DK transition (Figure 1E). As an example, *PLEC1*, encoding six isoforms of the keratinocyte adhesion protein plectrin, is transcribed from different promoters predicting KP-specific, DK-specific, and common isoforms (Figures 1F and S2B).

We annotated all CAGE promoters in six classes on the basis of the combinatorial presence of TATA box and CpG islands, i.e., $TATA^+$ or $TATA^-$, and no-CpG (NCPs), low-CpG (LCPs), and high-CpG (HCPs). The majority (~75%) of the promoters fell in the HCP class and were mostly $TATA^-$, a feature associated with housekeeping functions (Carninci et al., 2006; Schug et al., 2005). As expected, the proportion of LCP and NCP promoters progressively increased in the differentially expressed promoters at increasing FC values (Figure 1G).

Differential Promoter Usage Defines Stage-Specific Gene Expression Programs

Genes associated with differentially expressed CAGE promoters encoded known markers of follicular and interfollicular epidermal progenitors (i.e., *SOX9*, *LRIG1*, *BMI1*, *TCF3*, *TCF4*, *TP63*) and differentiating keratinocytes (*IVL*, *FLG*, *KRT1*, and genes belonging to the epidermal differentiation complex (EDC) on chromosome 1q21). To correlate differential promoter usage with gene-expression patterns, we carried out an RNA-sequencing (RNA-seq) analysis in KPs and DKs. The DK dataset showed a good correlation (Spearman's $r > 0.8$) with the RNA-seq data of human epidermal keratinocytes reported in ENCODE, demonstrating the similarity between the two populations and