

博士論文

ヒト角膜における細胞種特異的な遺伝子発現

およびクロマチン構造解析

中川 卓

## 目次

|   |    |
|---|----|
| 略語一覧 .....  | 3  |
| 「要旨」 .....  | 5  |
| I. 序文 .....   | 6  |
| 1A. 角膜の構造 .....   | 6  |
| 1B. 角膜輪部幹細胞 .....   | 6  |
| 1C. 角膜輪部上皮機能不全 .....  | 9  |
| 1D. 角膜輪部機能不全に対する角膜上皮の再生医療 .....                                 | 11 |
| 1E. 培養角膜輪部シートの培養法の進展 .....                                      | 15 |
| 1F. 角膜輪部機能不全の新たな治療 .....  | 16 |
| 1G. 角膜上皮のマーカー遺伝子と、それらを規定する転写因子 .....                            | 17 |
| II. 本研究の目的 .....  | 20 |
| III. 角膜上皮細胞の網羅的発現解析 .....                                       | 21 |
| 3A. 背景 .....  | 21 |
| 3B. 材料と方法 .....   | 22 |
| 3B1. 研究用ヒトドナー強角膜片 .....   | 22 |
| 3B2. ヒトドナー強角膜片からの角膜上皮細胞の採取と RNA の抽出 .....                       | 24 |
| 3B3. ヒト角膜輪部上皮細胞の初代培養 .....                                      | 26 |
| 3B4. ヒト培養細胞株 .....  | 28 |
| 3B5. Exon array を用いた網羅的な遺伝子発現解析 .....                           | 28 |
| 3B6. mRNA-seq を用いた網羅的な遺伝子発現解析 .....                             | 30 |
| 3B7. 逆転写、定量リアルタイム PCR .....                                     | 32 |
| 3B8. 免疫染色 .....   | 32 |
| 3C. 結果 .....  | 33 |
| 3C1. ドナー強角膜片より得られた RNA の量と質 .....                               | 33 |
| 3C2. Exon array を用いた、角膜上皮特異的なマーカー遺伝子の探索 .....                   | 35 |
| 3C3. mRNA-seq を用いた初代培養角膜輪部上皮細胞、角膜上皮細胞株<br>の遺伝子発現解析 .....        | 41 |
| 3C4. mRNA-seq、Exon array を用いた皮膚と比較して角膜上皮で発現が<br>高い転写因子の探索 ..... | 46 |
| IV. 角膜上皮細胞のクロマチン構造解析 .....                                      | 49 |

|   |    |
|---|----|
| 4A. 背景 .....  | 49 |
| 4A1. ヌクレオソーム・フリー領域 .....                                  | 49 |
| 4A2. FAIRE-seq 法 .....                                    | 50 |
| 4A3. モチーフ解析 .....   | 50 |
| 4B. 方法 .....  | 51 |
| 4B1. 概要 .....   | 51 |
| 4B2. FAIRE .....  | 51 |
| 4B3. FAIRE サンプルの濃縮確認 .....                                | 53 |
| 4B4. 次世代シーケンサーによる FAIRE サンプルの解析 .....                     | 54 |
| 4B5. シングルエンドライブラリー作成 .....                                | 54 |
| 4B6. ライブラリーの確認 .....                                      | 54 |
| 4B7. 次世代シーケンサーでの解析 .....                                  | 55 |
| 4B8. ヒトリファレンスゲノムへのマッピングとピークの抽出 .....                      | 55 |
| 4B9. 転写開始点からの位置によるヌクレオソーム・フリー領域の遺伝子ア<br>ノテーション .....      | 55 |
| 4B10. 公開データの利用による DNase-seq、CTCF ChIP-seq .....           | 56 |
| 4B11. 角膜上皮特異的エンハンサー領域の同定 .....                            | 58 |
| 4B12. モチーフ解析 .....  | 58 |
| 4C. 結果 .....  | 60 |
| 4C1. FAIRE-seq による角膜上皮細胞特異的なエンハンサー領域の同定                   | 60 |
| 4C2. 角膜上皮特異的エンハンサー領域のモチーフ解析 .....                         | 69 |
| 4C3. KRT12 近傍および KRT3 近傍の角膜上皮特異的エンハンサー領<br>域のモチーフ解析 ..... | 75 |
| V. 考察 .....   | 77 |
| 5A. 角膜上皮特異的遺伝子の同定 .....                                   | 77 |
| 5B. 初代培養角膜輪部上皮細胞と角膜上皮細胞株の遺伝子発現解析 .....                    | 78 |
| 5C. ヒトドナー角膜上皮細胞の FAIRE-seq .....                          | 80 |
| 5D. KRT3、KRT12 近傍のエンハンサー領域の同定 .....                       | 80 |
| 5E. エンハンサー領域と遺伝子発現の相関 .....                               | 81 |
| 5F. 角膜上皮細胞の特異性維持に関わる 7 転写因子 .....                         | 82 |
| 5G. 今後の展開 .....   | 84 |
| VI. まとめ .....   | 87 |
| 「謝辞」 .....  | 89 |
| 「引用文献」 .....  | 91 |

## 略語一覧

CEpi: human donor corneal epithelial cells

CEpi cell line: SV40 large T antigen immortalized human corneal epithelial cell line

PrimaryLEpi: primary cultivated human limbal epithelial cells

ARPE19: human retinal pigmented epithelial cell line

LensEpi: human lens epithelial cell line

NHEK: normal human epidermal keratinocytes

FAIRE: Formaldehyde-Assisted Isolation of Regulatory Elements

Exon array: Affymetrix GeneChip Human Exon 1.0 ST Array

RNA: ribonucleic acid

DNA: deoxyribonucleic acid

mRNA-seq: mRNA-sequencing

mRNA: messenger RNA

FPKM: fragments Per Kilobase of exon per Million mapped fragments

TSS: transcription start sites

TFs: transcription factors

RT-PCR: Reverse Transcription Polymerase Chain Reaction

TRANSFAC: 転写因子とその認識配列情報に関するデータベース

JASPAR: 転写因子とその認識配列情報に関するデータベース

MODIC: Motif Identification Algorithm through Direct Comparison of Signal/Noise

Distributions Based on Maximum Entropy Method

TRAP: transcription factor binding affinity prediction

Genomatix: genomatix software suite

ENCODE: The Encyclopedia of DNA Elements

FBS: fetal bovine serum

BPE: bovine pituitary extract

KRT3: keratin 3

KRT12: keratin 12

PAX6: paired box 6

FOSL1: fos-like antigen 1

NR1D1: nuclear receptor subfamily 1, group D, member 1

SOX15: SRY (sex determining region Y)-box 15

FOXP2: forkhead box P2

HLF: hepatic leukemia factor

MEIS1: Meis homeobox 1

KLF3: Kruppel-like factor 3

KLF7: Kruppel-like factor 7

KLF10: Kruppel-like factor10

TGIF1: TGFB-induced factor homeobox 1

BHLHE41: basic helix-loop-helix family, member e41

SMAD3: SMAD family member 3

TP63: tumor protein p63

ABCG2: ATP-binding cassette, sub-family G (WHITE), member 2

CDH2: cadherin 2, type 1, N-cadherin (neuronal)

KRT19: keratin 19

NGFR: nerve growth factor receptor

ITGA6: integrin, alpha 6

## 「要旨」

角膜上皮細胞は、非角化重層扁平上皮であり、特異的なマーカーとして **KRT3** と **KRT12** が報告されているが、このような角膜上皮細胞の特異性を規定する転写因子は明らかでない。今回、ヒト角膜上皮細胞の網羅的な遺伝子発現解析およびクロマチン構造解析を行い、角膜上皮細胞の特異的遺伝子とそれらを規定する転写因子の同定を試みた。まず、Exon array にて 81 個のヒト正常組織・培養細胞と比較して角膜上皮細胞(組織)で特異的に発現しているマーカー遺伝子候補を 52 個見出した。次に、無血清・無フィーダー培養による初代培養角膜輪部上皮細胞と、角膜上皮細胞株における 52 個のマーカー遺伝子の発現を mRNA-seq を用いて探索した。さらに、Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)-sequencing 法を用いて角膜上皮細胞のヌクレオソーム・フリー領域を 35,952 箇所同定し、ヒト 9 細胞と比較することにより 2,622 箇所の角膜上皮特異的エンハンサー領域を同定した。角膜上皮特異的マーカー **KRT12**、**KRT3** の近傍にも角膜上皮特異的エンハンサー領域が存在した。次に、2,622 箇所のエンハンサー領域に濃縮している特徴的な転写因子の結合配列(モチーフ)を解析したところ、モチーフが濃縮しておりかつ角膜上皮細胞で発現が上昇している転写因子が 13 個確認された。このうち 7 転写因子は **KRT12** 近傍のエンハンサー領域にモチーフが確認され、角膜上皮の特異性を規定する可能性が高い転写因子と考えられた。

## I. 序文

角膜上皮の特異性を規定する転写因子を明らかにするために、ヒト角膜上皮細胞の網羅的な遺伝子発現解析およびクロマチン構造解析を行った。まず、背景と関連する既報研究について概観する。

### 1A. 角膜の構造

角膜は眼球の最前面に位置し、無血管で透明な組織である（図 1A）。強膜と共に眼球の外層を形成しており、眼球の内容を守るバリア機能を果たしている。また、角膜は光を通す透明な窓としての役割があり、強膜は網膜に像を形成するための暗箱を提供する。角膜は常に外部環境にさらされ続けるが、強膜は半透明の結膜により覆われており、直接外部にはさらされていない。

角膜は約 500  $\mu\text{m}$  の厚みがある。角膜は、非角化重層扁平上皮である角膜上皮、中層の結合組織である角膜実質、最内層の角膜内皮の三層構造である（図 1B）。三層のどの層が障害されても、角膜混濁と視力障害を来す。

### 1B. 角膜輪部幹細胞

角膜上皮は、5~7 層に重層化しており、基底細胞、翼状細胞、最表層細胞から成る（図 1C）。角膜上皮はタイトなバリア機能を有し、外界から眼球を守る役割を果たしている。この角膜上皮の幹細胞は角膜と結膜の境界部の角膜輪部と呼ばれる特別なニッチの基底層にあると考えられている（角膜輪部幹細胞）（図 1D）<sup>1-3</sup>。

角膜輪部幹細胞により一過性増殖細胞 (Transient amplifying (TA) 細胞) が作られ、中央角膜に移動していく。角膜輪部幹細胞は slow-cycling と考えられているが、TA 細胞は急速に増殖する。角膜上皮は角膜輪部幹細胞により絶えず更新され、維持されていると、推測されている (図 1D) <sup>3,4</sup>。

現在までに、TP63 や、ABCG2、CDH2、KRT19、NGFR、ITGA6 が角膜輪部幹細胞のマーカ候補として報告されている <sup>4-8</sup> が、統一的な見解は得られていない。



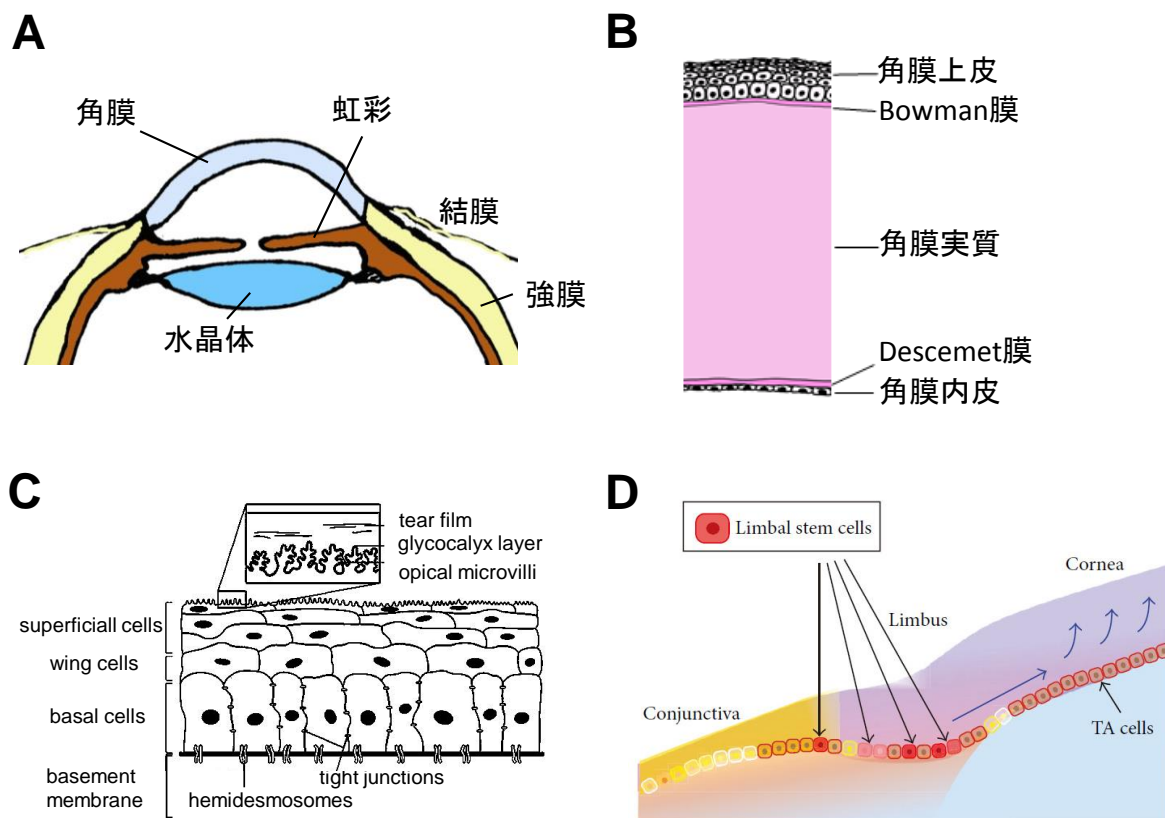


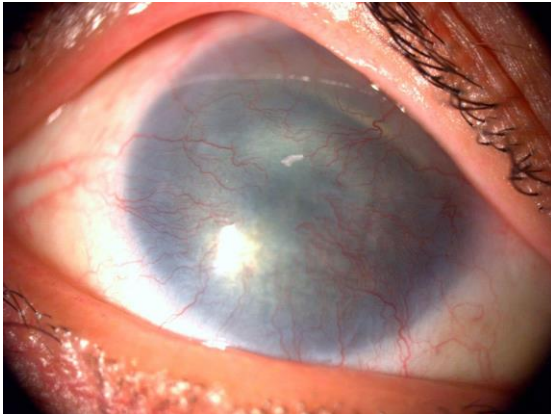
図 1 : 角膜の構造

(A) 前眼部の解剖図。角膜は眼球の最前面に位置する。(B)ヒト角膜の模式図。角膜は上皮、実質、内皮の三層構造をとる。(C)ヒト角膜上皮の模式図。角膜上皮細胞は非角化重層扁平上皮であり、小型の基底細胞、中型の翼状細胞、多角形の扁平表層細胞から成る。(D)角膜輪部幹細胞(Limbal stem cells)の模式図。角膜輪部幹細胞は角膜輪部(Limbus)の基底部に散在する。一過性増殖細胞(TA細胞)は角膜輪部幹細胞より分化した前駆細胞であり、中央角膜に移動する。(C:文献 Farjo A et al. Ophthalmology 3<sup>rd</sup> ed. Mosby、2008 より一部改変、D: 文献<sup>3</sup>より一部改変)。

### 1C. 角膜輪部上皮機能不全

角膜疾患を原因とする失明は世界で2番目に多く、全世界で600-800万人程いると推定されている<sup>9</sup>。そのうちいくつかの角膜疾患には薬剤や手術により治療できるが<sup>10-12</sup>、角膜輪部幹細胞が障害される角膜輪部機能不全では従来型の全層角膜移植では治療が困難である。角膜輪部機能不全の原因としては、アルカリや酸による化学熱傷、あるいはStevens-Johnson症候群、眼類天疱瘡のような炎症性・免疫性疾患、無虹彩症のような先天疾患などがある。角膜輪部に存在する角膜輪部幹細胞が障害されると角膜上皮細胞が供給されなくなり、結膜上皮が角膜表面を覆い角膜の透明性は失われ失明に至る<sup>3,13-15</sup> (図2)。

**A**



**B**

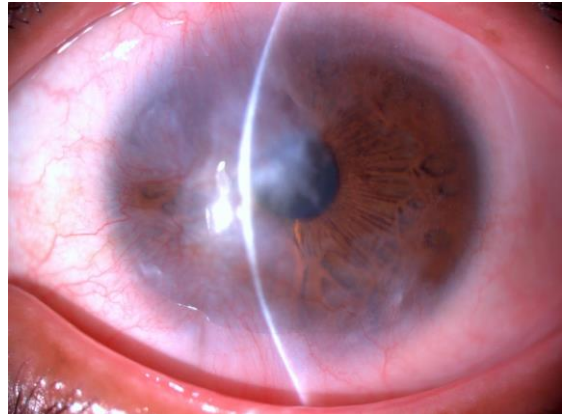


図 2 : 角膜輪部機能不全患者の症例写真。

(A) 無虹彩症。(B) セメント(アルカリ)外傷。角膜輪部の幹細胞が障害される角膜輪部機能不全では、周囲からの結膜線維組織侵入・角膜血管新生・角膜混濁による著明な視力障害を認める。

## 1D. 角膜輪部機能不全に対する角膜上皮の再生医療

前述のように角膜輪部機能不全は治療が非常に困難である。角膜輪部機能不全に対して、角膜内皮や角膜実質の障害に対して通常行われる従来型の全層角膜移植では、一時的な角膜透明性の改善は得られるものの、数か月で結膜上皮が角膜表面を覆ってしまう。角膜上皮の幹細胞が角膜輪部に存在することが分かってくると1990年ごろより同種の角膜輪部移植が行われたが、高率な拒絶反応やドナー不足が問題として残った<sup>16</sup>。一方、1989年Kenyonらにより自己角膜輪部を移植するモデルが提唱された<sup>17</sup>。しかし自己移植の場合、正常な角膜輪部が片眼に十分に残されている必要があり、両眼性の疾患には適応できない。

1997年にPellegriniらにより患者自身の角膜輪部を培養して角膜上皮シートを作製する手法で低侵襲での角膜輪部機能不全に対しての治療が可能となった<sup>18</sup>。両眼とも障害された疾患に対処するために2003年にNakamuraらが<sup>19</sup>、2004年にはNishidaらが<sup>20</sup>、培養した自己口腔粘膜上皮を眼表面に移植することで角膜輪部機能不全を治療できる可能性を報告した<sup>21</sup>。

しかし、培養口腔粘膜上皮移植を用いても、特にStevens-Johnson症候群のような炎症性眼表面疾患では依然として難治であり、①移植後の角膜周辺からの新生血管侵入、②上皮欠損やそれに伴う感染症、といった問題点が未解決である<sup>22,23</sup>。また、2006年、Taniokaらにより、自己結膜を利用した培養自己結膜上皮移植も報告された<sup>24</sup>（図3、図4）。しかし、結膜移植は、健常な結膜が利用可能な、片眼性の疾患または非炎症性の両眼性の疾患に限られる<sup>25</sup>（図

4) 。また、他の自家細胞源としては、iPS 細胞<sup>14,26</sup>、歯髄幹細胞、毛包幹細胞、骨髄由来間葉系幹細胞等が動物実験レベルで試みられている<sup>27</sup>。しかし、ES 細胞<sup>28</sup>、iPS 細胞<sup>14,26</sup>からの角膜上皮細胞への分化誘導に関しては、純度や分化効率の問題もあり、分化誘導法の確立には至っていない。

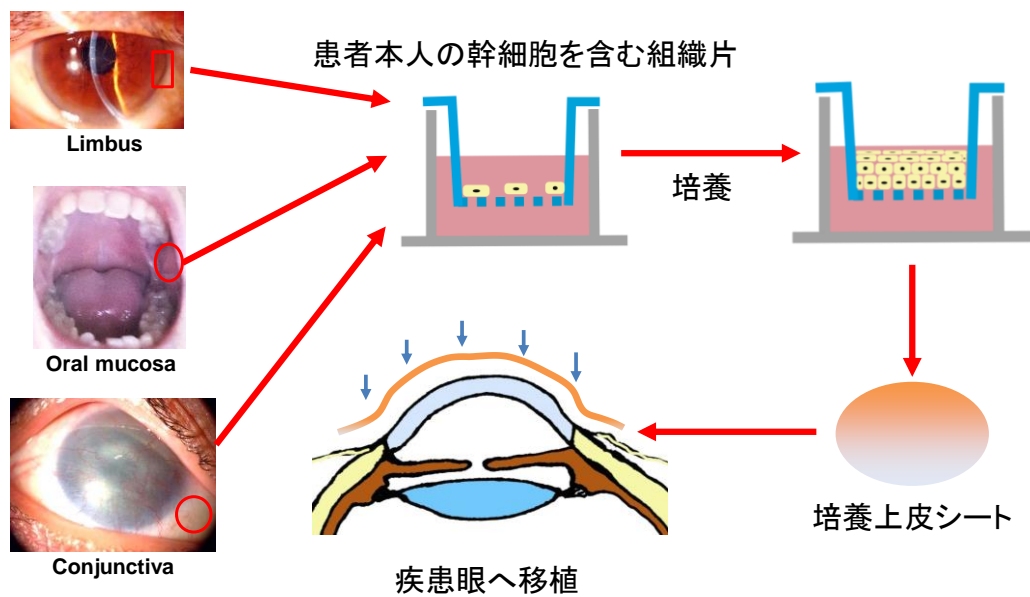


図 3 : 培養自己上皮幹細胞シート移植による角膜輪部機能不全の治療

自己の角膜輪部 (Limbus)、口腔粘膜 (Oral mucosa) または結膜 (Conjunctiva) を採取し、3 週間前後の培養し、培養重層化上皮シートを作成する。培養シートを疾患眼へ移植し、眼表面を再建する。

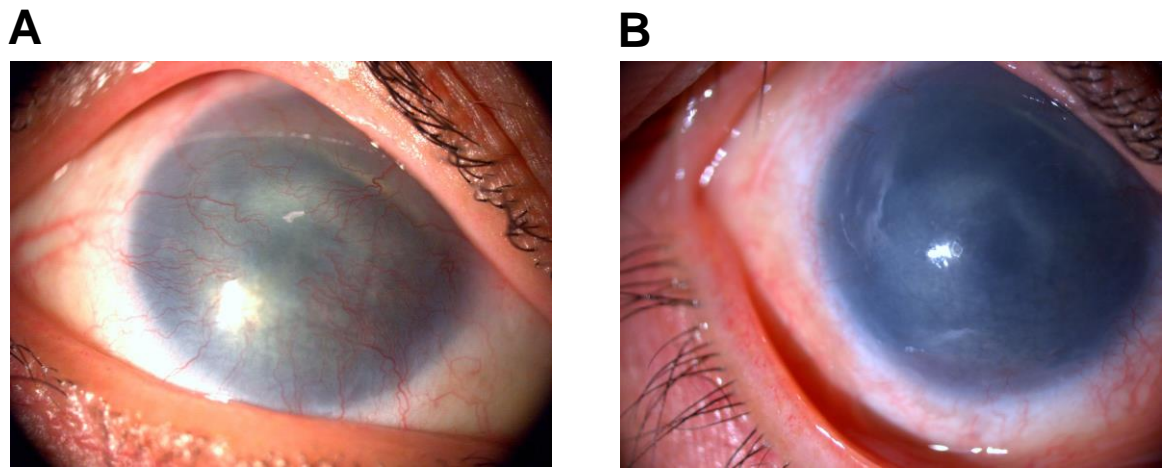


図 4 : 培養自己結膜上皮シート移植前後の角膜輪部機能不全患者の症例写真。  
(A) 無虹彩症による角膜輪部機能不全(図 2A に同じ)。(B) 培養自己結膜上皮シート移植後 3 か月後。術前に比べて、角膜新生血管は抑えられ、角膜の透明性は改善している。

## 1E. 培養角膜輪部シートの培養法の進展

前述したように、1997年に Pellegrini らにより患者自身の角膜輪部を培養して角膜上皮シートを作製する手法で低侵襲での難治性角膜上皮疾患克服がある程度は可能となった<sup>18</sup>。しかし、培養にはマウス 3T3 細胞（フィーダー細胞）、胎仔血清（FBS）含有培地、ウシ下垂体抽出物（BPE）が用いられ、動物由来の未知の成分による潜在的な汚染が、懸念された<sup>29,30</sup>。それに対して、2008年、我々はフィーダー細胞、FBS、BPE をいずれも用いない培養法として、B27 supplement を用いた無血清・無フィーダーの角膜上皮シート作成に成功した<sup>31,32</sup>。同方法で作成した培養シートは、生体内と同様に4～5層に重層化しており、角膜上皮マーカーである KRT3、KRT12 を発現していた。また、カルボキシフルオレセインの透過性を測定することでバリア機能の測定も可能であり、臨床的に眼表面の再建に用いるのみならず、生体内で4～5層の重層非角化扁平上皮を構成する角膜上皮を再現した *in vitro* モデルとしての可能性が示された<sup>32</sup>。



## 1F. 角膜輪部機能不全の新たな治療

これまでに述べたような背景を踏まえ、治療が困難な角膜輪部機能不全に対する新たな治療として、以下の2つを考えるに至った。

1. 表皮ケラチノサイトに角膜上皮の特異性を規定する転写因子を導入することにより、*in vitro*での角膜上皮細胞へのダイレクトリプログラミングによる、遺伝的に改良した培養上皮シートを用いた治療。
2. 角膜輪部機能不全患者に角膜上皮特異性を規定する転写因子を導入することにより、*in vivo*での角膜上皮細胞へのダイレクトリプログラミングによる、角膜輪部機能不全患者の治療。

1. に関しては、角膜上皮の理想的な再生医療として、障害されていない部位から容易に細胞源を採取できること、ダイレクトリプログラミングによりできるだけ角膜上皮に似た遺伝子プロファイルの重層化した上皮シートが得られることが挙げられる。

2.に関しては、角膜は最表層に位置し、薬剤徐放性のコンタクトレンズや点眼を用いるなどすれば、*in vivo*での遺伝子導入のハードルが他組織に比べて容易であること、ダイレクトリプログラミングにより新生血管の退縮、眼表面のバリア機能の改善や角膜透明性の向上が期待できることが挙げられる。

そのためには、角膜上皮の特異性を規定する転写因子を明らかにすることが必要である。

例えば、これまで、肝細胞の転写因子ネットワークが明らかにされてきており<sup>33,34</sup>、そのような知見を背景として、2011年に Suzuki らが肝細胞で重要と考

えられてきた転写因子を線維芽細胞に導入することで肝細胞へのダイレクトプログラミングに成功している<sup>35</sup>。

## 1G. 角膜上皮のマーカー遺伝子と、それらを規定する転写因子

角膜上皮細胞は特異的マーカーとしてケラチン 3 (KRT3) とケラチン 12 (KRT12) が知られており<sup>1,36,37</sup>、角膜上皮への分化や誘導の指標となっている。ケラチンは、上皮細胞において、直径 10 nm の中間径フィラメントを構成する非水溶性のタンパク質であり、機械的ストレスから上皮細胞を安定させる役割がある。50 以上のケラチンが知られており<sup>38-40</sup>、酸性のタイプ I ケラチンと塩基性のタイプ II ケラチンに分けられる。生体内では、塩基性ケラチンは特定の酸性ケラチンとペアとなり、ヘテロポリマーを形成し共発現している。ケラチンペアの発現は組織特異的、分化依存的に制御されている。KRT3/KRT12 ペアの発現は、ヒトやウシ、ブタ、ウサギ、ニワトリの角膜上皮の分化マーカーとされている<sup>1,36</sup>。KRT12 の発現は角膜上皮に限定されていると報告されている<sup>36,41</sup>。しかし、KRT3 は角膜上皮以外の組織での発現も報告されており<sup>36,42</sup>、マウス角膜上皮では KRT3 の発現が確認されていない<sup>43,44</sup>。また、KRT3、KRT12 は Meesmann 角膜上皮ジストロフィの原因遺伝子であり<sup>45</sup>、Krt12 のノックアウトマウスは角膜上皮の脆弱性を来すことが報告されている<sup>46</sup>。角膜上皮の表現型を保持する確立した細胞株は現在のところ知られていないため、*in vitro* での細胞数を要する角膜上皮の特異性の研究は難しいという問題がある。

これまで、角膜上皮の特異性を規定する転写因子に関連して、**PAX6**を導入すると、**KRT3**、**KRT12** の発現が上昇するとの報告や<sup>47</sup>、角膜輪部機能不全において、**PAX6**と**KRT3**、**KRT12** の発現低下の関連を示唆する報告がみられる<sup>48</sup>。本研究では、ヒト角膜上皮細胞における網羅的な発現解析と **FAIRE-seq** によるクロマチン構造解析を行うことにより、角膜上皮細胞の特異性を規定する転写因子の同定を目的とした（図5）。

## 角膜上皮細胞

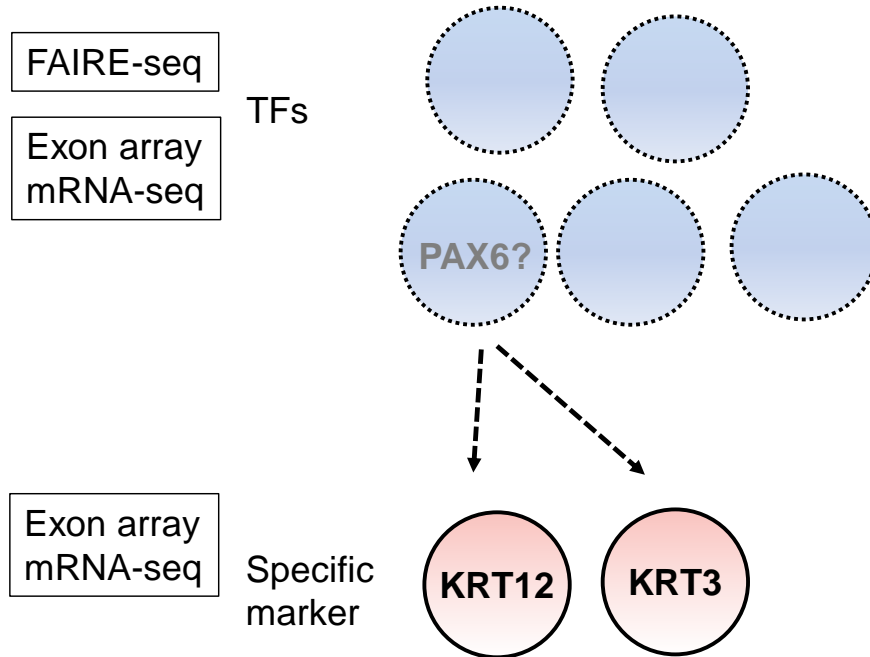


図 5: ヒト角膜上皮細胞の特異性を規定する転写因子の同定

網羅的発現解析 (Exon array、mRNA-seq) により、角膜上皮特異的マーカー遺伝子を同定する。FAIRE-seq によるクロマチン構造解析と網羅的発現解析の統合解析により、角膜上皮の特異性を規定する転写因子 (TFs) を同定する。

## II. 本研究の目的

角膜上皮細胞の特異性を規定する転写因子を同定することを目的とした。

上記の目的に従い、研究は以下のように行った。

1. Exon array を用いた網羅的発現解析により、角膜上皮細胞（組織）の特異的マーカー遺伝子をゲノムワイドに探索する。
2. SV40 large T 抗原不死化ヒト角膜上皮細胞株、初代培養角膜輪部上皮細胞における角膜上皮の特異的マーカー遺伝子の発現を mRNA-seq、定量 RT-PCR にて確認する。
3. mRNA-seq と Exon array を用いた網羅的発現解析により、角膜上皮で発現が上昇している転写因子を探索する。
4. FAIRE-seq により、角膜上皮の特異的エンハンサー領域を明らかにし、3.で求めた角膜高発現転写因子の中で、角膜上皮特異的エンハンサー領域にモチーフが濃縮している転写因子を同定する。
5. 角膜上皮の特異的マーカー遺伝子の近傍のエンハンサー領域について、4.で同定した転写因子のモチーフの配列解析を行う。

以下、2 つに分けて叙述する。

角膜上皮細胞の網羅的発現解析

角膜上皮細胞のクロマチン構造解析

### Ⅲ. 角膜上皮細胞の網羅的発現解析

#### 3A. 背景

網羅的発現解析は、Exon arrayとmRNA-seqを用いた。Human Exon 1.0 ST array (Affymetrix) は東京大学先端科学技術研究センター ゲノムサイエンス分野 油谷浩幸 研究室のヒトサンプルデータ (ヒト正常組織・培養細胞81サンプル) との比較が可能である。また、mRNA-seqは、microarrayに比べてより低発現領域の遺伝子発現の正確性と定量性が高いことが知られている<sup>49,50</sup>。そのため、まず、油谷研の多数のヒトサンプルと比較できるHuman Exon 1.0 ST arrayを用いて角膜上皮の特異的マーカー遺伝子の検索を行った。次に、低発現遺伝子に正確性の高いmRNA-seqを重視して皮膚と比較して高発現の転写因子の同定を試みた。

### **3B. 材料と方法**

#### **3B1. 研究用ヒトドナー強角膜片**

研究用ヒトドナー強角膜片は、アメリカアイバンク (SightLife) より輸入した。9ドナーの両眼(18眼)と1ドナーの1眼、合計19眼を用いた。ドナーの年齢は57-75歳、人種は全てCaucasianで、内眼手術の既往が無く、感染症(HIV、HBV、HCV、HTLV)が陰性のものを用いた(表1)。死後10時間以内に摘出され、4℃にて保存液(Optisol corneal storage medium、Chiron Ophthalmics)に保存された。全ての検体で死後10日以内にRNA (mRNA-seq、Exon array、qRT-PCR)またはDNA (FAIRE-seq)を抽出した。

表 1 : 本研究に用いた研究用ヒトドナー強角膜片のプロファイル。

| Tissue # | Age | sex    | cause of death           | RNA-seq_1 | RNA-seq_2 | Exon array_1 | Exon array_2 | qRT-PCR | FAIRE_1 | FAIRE_2 |
|----------|-----|--------|--------------------------|-----------|-----------|--------------|--------------|---------|---------|---------|
| 1        | 57  | Male   | Pancreatic cancer        | •         |           |              |              |         |         |         |
| 2        | 57  | Male   | Pancreatic cancer        | •         |           |              |              |         |         |         |
| 3        | 64  | Male   | Cholangiocarcinoma       |           |           |              |              | •       |         |         |
| 4        | 64  | Male   | Cholangiocarcinoma       |           |           |              |              | •       |         |         |
| 5        | 73  | Male   | Cancer                   |           |           | •            |              |         |         |         |
| 6        | 73  | Male   | Cancer                   |           |           | •            |              |         |         |         |
| 7        | 74  | Female | COPD                     |           | •         |              |              |         |         |         |
| 8        | 74  | Female | COPD                     |           | •         |              |              |         |         |         |
| 9        | 75  | Female | Cancer                   |           | •         |              |              |         |         |         |
| 10       | 75  | Female | Cancer                   |           | •         |              |              |         |         |         |
| 11       | 69  | Male   | Cancer                   |           |           |              | •            |         |         |         |
| 12       | 69  | Male   | Cancer                   |           |           |              | •            |         |         |         |
| 13       | 66  | Female | Uterine cancer           |           |           |              | •            |         |         |         |
| 14       | 66  | Female | Uterine cancer           |           |           |              | •            |         |         |         |
| 15       | 69  | Female | CVA                      |           |           |              |              |         | •       |         |
| 16       | 69  | Female | CVA                      |           |           |              |              |         | •       |         |
| 17       | 67  | Female | ICB/ ICH                 |           |           |              |              |         | •       |         |
| 18       | 67  | Female | ICB/ ICH                 |           |           |              |              |         | •       |         |
| 19       | 65  | Female | Metastatic Breast cancer |           |           |              |              |         |         | •       |



### 3B2. ヒトドナー強角膜片からの角膜上皮細胞の採取とRNAの抽出

ヒトドナー強角膜片に由来する角膜上皮細胞は、実体顕微鏡 (SZX12、Olympus) 下で氷上の PBS 溶液 (NaCl: 137 mM、KCl: 2.68 mM、Na<sub>2</sub>HPO<sub>4</sub>: 8.1 mM、KH<sub>2</sub>PO<sub>4</sub>: 1.47 mM、pH 7.4) 中で機械的に剥離し採取した (図 6)。角膜上皮細胞は、隣接した角膜実質の混入無く、Bowman 膜上から滑らかに採取された。

採取した角膜上皮細胞は、直ちに RTL buffer 350 µl に溶解し、RNeasy microkit (Qiagen) により RNA を抽出した。純水 14 µl をカラムに添加し、RNA を溶出した。その後、分光光度計 (Nanodrop ND-1000 Spectrophotometer、Thermo Fisher Scientific) により RNA の濃度を測定した。また、バイオアナライザー (Agilent) により RNA の質と濃度を測定した。精製した RNA は -80°C で保存した。

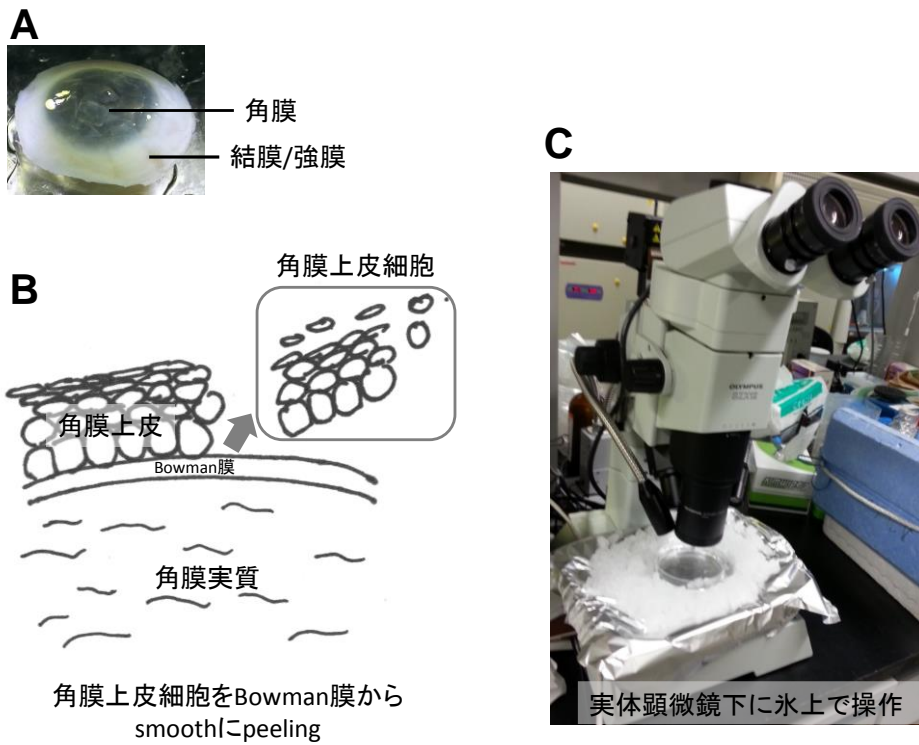


図 6: ヒト角膜上皮細胞の採取

(A) 研究用ヒトドナー強角膜片 (B) 研究用ヒトドナー強角膜片より角膜上皮をBowman膜から機械的に剥離した。(C) 実体顕微鏡 SZX12 (Olympus)。

### 3B3. ヒト角膜輪部上皮細胞の初代培養

ヒト初代培養角膜上皮細胞の培養は、既報<sup>31,32</sup>を改変した方法で行った。具体的には、研究用ヒトドナー強角膜片(3B1 参照)より、結膜を剪刃にて除去後、角膜輪部組織を剪刃にて採取した。次に 0.02% タイプ 1A コラゲナーゼ(Wako)を含む DMEM/F12 培地中で 37°C、5% CO<sub>2</sub>にて一晩静置した。翌日、0.05% Trypsin/EDTA 中で 37°Cにて 10 分間静置した後ピペッティングを行い、ポアサイズ 0.45 μm のセルストレーナー (Falcon) を通過させて残存組織を除去した。2.5×10<sup>4</sup> cells/cm<sup>2</sup>にて細胞をカルチャーインサート上に播種した。培地は、DMEM/HamF12(1:1)、B27 supplement (Invitrogen)、epidermal growth factor (EGF) 20 ng/ml を用い、無血清・無フィーダーで培養を行った。1 日おきに培地交換し、10 日前後で confluent となり、その後は毎日培地交換を行った。約 20 日間で、4~5 層に重層化した上皮シートを得た(図 7)。

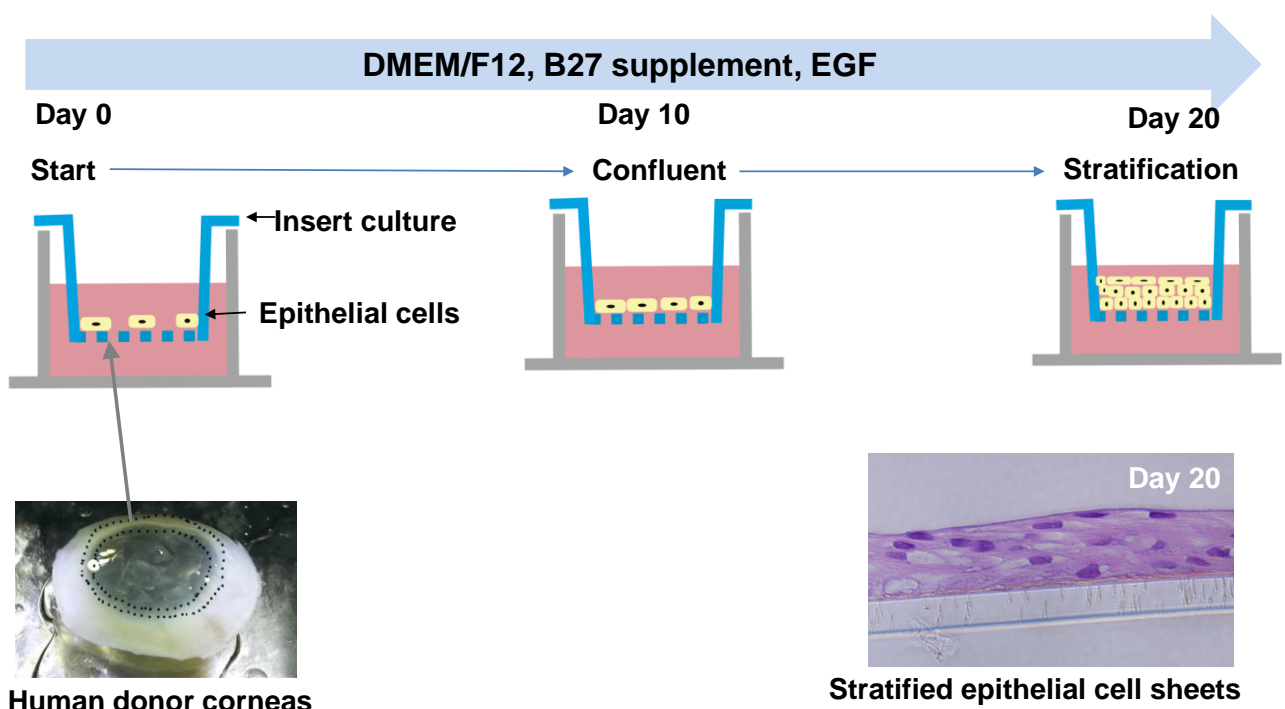


図 7: ヒト角膜輪部上皮細胞の初代培養

研究用ヒトドナー強角膜片より角膜輪部上皮を採取し、細胞を播種した。約 3 週間で重層化した上皮シートを得た。

### 3B4. ヒト培養細胞株

角膜上皮細胞株は、Araki-Sasaki らがヒト初代培養角膜上皮細胞を SV40 large T 抗原にて不死化した角膜上皮細胞株を用いた<sup>51</sup>。培養は、DMEM/ F12 FBS 10%、37°C、5% CO<sub>2</sub> 中で行った。また、ヒト角膜輪部上皮細胞の初代培養と同様の無血清・無フィーダーの培養条件(前項参照)でも培養を行った。また、比較対象として、網膜色素上皮細胞株 (ARPE19) (培地:DMEM/ F12 FBS 10%、継代)、SV40 largeT 抗原不死化ヒト水晶体上皮細胞株 (LensEpi) (培地:DMEM、FBS 20%)、正常ヒト表皮角化細胞 (HEKN) (培地:Keratinocyte-SFM、Invitrogen)、正常ヒト皮膚線維芽細胞 (成人) (NHDF) (Gibco) (培地:DMEM FBS 10% または無血清培地: DMEM/F12、B27 supplement、bFGF 20 ng/ ml)を準備し、凍結保存されている細胞を 2 回継代した後を用いた。いずれも 37°C、5% CO<sub>2</sub> で培養を行った。培地は 1 日おきに交換した。

### 3B5. Exon arrayを用いた網羅的な遺伝子発現解析

Human Exon 1.0 ST Array (Affymetrix) を用いて、ヒトドナー角膜上皮細胞 (CEpi: 2 サンプル) の転写産物を網羅的に解析した。比較対象として、油谷研にてすでに取得していたヒト正常組織および培養細胞の合計 81 サンプルの発現データを用いた (表 2)。遺伝子ごとに各 Core プローブのシグナル値の幾何平均を計算し、17,299 個の各遺伝子の発現値 (Raw signal) とした。

表 2 : 本研究で用いた Exon array の材料(A)と、比較対象として用いたヒト正常組  
織・培養細胞 81 サンプルの Exon array のデータセット(B)

(A)

| Sample name | Description                         | Race      |
|-------------|-------------------------------------|-----------|
| CEpi1       | Human donor corneal epithelial cell | Caucasian |
| CEpi2       | Human donor corneal epithelial cell | Caucasian |

(B)

| Sample name           | Description                         | Race              |
|-----------------------|-------------------------------------|-------------------|
| tongue                | tongue                              | Unknown           |
| esophagus             | esophagus                           | Caucasian         |
| esophagus_2           | esophagus                           | Unknown           |
| tp_trachea            | trachea                             | Caucasian         |
| trachea_2             | trachea                             | Unknown/Caucasian |
| saec                  | Human Small Airway Epithelial Cells |                   |
| HMEC_5F1606           | Human Mammary Epithelial Cells      |                   |
| HMEC_6F340B           | Human Mammary Epithelial Cells      |                   |
| cervix_2              | Cervix                              | Unknown/Caucasian |
| skin                  | Skin                                | Unknown           |
| brain                 | brain                               | Asian             |
| cortex                | cortex                              | Unknown           |
| pons                  | pons                                | Unknown           |
| hippocampus           | hippocampus                         | Unknown           |
| diencephalon          | diencephalon                        | Unknown           |
| thalamus              | thalamus                            | Unknown           |
| cerebellum            | cerebellum                          | Asian             |
| oblongata             | oblongata                           | Unknown           |
| pituitary             | pituitary                           | Caucasian         |
| dorsal_root_ganglion  | dorsal_root_ganglion                | Caucasian         |
| olfactory             | olfactory                           | Unknown           |
| tp_salivarygland      | salivary gland                      | Caucasian         |
| tonsil                | tonsil                              | Caucasian         |
| thyroid               | thyroid                             | Unknown           |
| s_muscle              | smooth muscle                       | Caucasian         |
| heart                 | heart                               | Caucasian         |
| heart_2               | heart                               | Unknown           |
| lungadeno2n           | lungadeno                           | Unknown           |
| lungadeno22n          | lungadeno                           | Unknown           |
| normallung_clontech   | normal lung_clontech                | Caucasian         |
| normallung_ambion     | normal lung_ambion                  | Caucasian         |
| normallung_stratagene | normal lung_stratagene              | Unknown           |
| normallung_panomics   | normal lung_panomics                | Unknown           |
| lung                  | lung                                | Japanese          |
| stomach               | stomach                             | Japanese          |
| intestine             | intestine                           | Caucasian         |
| colon                 | colon                               | Caucasian         |
| pancreas              | pancreas                            | Caucasian         |
| pancreas_2            | pancreas                            | Unknown           |
| pancreas_3            | pancreas                            | Caucasian         |
| nl_1                  | liver                               | Unknown/Caucasian |
| nl_2                  | liver                               | Unknown           |
| liver_2               | liver_2                             | Caucasian         |
| liver_3               | liver_3                             | Caucasian         |
| gallbladder           | gallbladder                         | Unknown           |
| tp_adrenalgland       | adrenal gland                       | Caucasian         |
| kidney                | kidney                              | Caucasian         |
| kidney_2              | kidney                              | Unknown           |
| kidney_3              | kidney                              | Caucasian         |
| tp_bladder            | bladder                             | Caucasian         |
| spleen                | spleen                              | Caucasian/Black   |
| adipose               | adipose                             | Unknown           |
| artery                | artery                              | Unknown           |
| vein                  | vein                                | Unknown           |
| lymphnode             | lymphnode                           | Unknown           |
| lymphnode_2           | lymphnode                           | Unknown           |
| bonemarrow            | bonemarrow                          | Caucasian         |
| peripheralblood       | peripheralblood                     | Unknown           |
| tp_macrophage         | macrophage                          | Unknown           |
| tp_monocyte           | monocyte                            | Unknown           |
| pericardium           | pericardium                         | Caucasian         |
| cavalvein             | cavalvein                           | Unknown           |
| breast                | breast                              | Unknown           |
| breast24n             | breast                              | Unknown           |
| ovary                 | ovary                               | Caucasian         |
| tp_uterus             | uterus                              | Unknown           |
| placenta              | placenta                            | Black             |
| prostate              | prostate                            | Caucasian         |
| prostate_gsm310793    | prostate_gsm310793                  | Unknown           |
| prostate_gsm310794    | prostate_gsm310794                  | Unknown           |
| prostate_gsm310795    | prostate_gsm310795                  | Unknown           |
| testis                | testis                              | Caucasian         |
| seminalvehicle        | seminalvehicle                      | Unknown           |
| fetalbrain            | fetalbrain                          | Caucasian         |
| fetalcolon            | fetalcolon                          | Unknown           |
| fetaliver             | fetaliver                           | Unknown           |
| fetalilung            | fetalilung                          | Unknown           |
| tp_fetalstomach       | tp_fetalstomach                     | Unknown           |
| KhES1                 | human embryonic stem cells          |                   |
| KhES2                 | human embryonic stem cells          |                   |
| KhES3                 | human embryonic stem cells          |                   |

### 3B6. mRNA-seqを用いた網羅的な遺伝子発現解析

Illumina 社の次世代シーケンサー (HiSeq2000) を用いた mRNA-seq により、ヒトドナー角膜上皮細胞 (CEpi: 2 サンプル) の転写産物を網羅的に解析した。TruSeq RNA Sample Preparation Kits v2 (Illumina) を用いて、ライブラリ作成を行った。HiSeq 2000 を用い、100 bp、ペアードエンドのリードを得た。BWA aligner を用いてリファレンスゲノム (Human、GRC37、HG19) へのマッピングを行った。Refseq に記載されている遺伝子と John Rinn<sup>52</sup>らのまとめた長鎖非コード RNA (lncRNA) 等を合わせた計 86,187 個の転写産物について Fragments Per Kilobase of exon per Million mapped fragments (FPKM) を算出した。

同様に、ヒト初代培養角膜輪部上皮細胞 (PrimaryLEpi)、SV40 large T 不死化角膜上皮細胞株 (CEpi cell line)、網膜色素上皮細胞株 (ARPE19)、水晶体上皮細胞株 (LensEpi)、正常ヒト表皮角化細胞 (HEKN)、正常ヒト皮膚線維芽細胞 (NHDF)、無血清培養 NHDF (NHDF FBS (-)) の mRNA-seq も施行した (表 3)。

油谷研 Body map プロジェクトから、ヒト成人組織 15 種類、胎児組織 5 種類、国際エピゲノムコンソーシアムから血管内皮細胞 15 サンプルのデータを利用した。ライブラリ作成、マッピング、FPKM の算出は同様に行った (表 3)。

また、ENCODE プロジェクト<sup>53</sup> の mRNA-seq の以下の 7 細胞 (リンパ芽球細胞株 (GM12878)、ヒト胚性幹細胞 (H1-ES)、子宮頸癌細胞株 (HeLa-S3)、肝細胞癌細胞株 (HepG2)、ヒト臍帯静脈内皮細胞 (HUVEC)、慢性骨髄性白血病細胞株 (K562)、正常ヒト表皮角化細胞 (NHEK) (Duke 大学) の bam ファイルデータを、角膜上皮細胞と同様の発現解析パイプラインを利用して FPKM を算出した (表 3)。

表 3 :本研究で用いた mRNA-seq の材料(A)と、比較対象として用いたヒト正常組織・培養細胞 42 サンプルの mRNA-seq のデータセット(B)

(A)

| Sample name    | Description                                | Race      |
|----------------|--|-----------|
| CEpi1          | Human donor corneal epithelial cell        | Caucasian |
| CEpi2          | Human donor corneal epithelial cell        | Caucasian |
| PrimaryLEpi    | Primary cultivated limbal epithelial cell  | Caucasian |
| CEpi cell line | Human corneal epithelial cell line         |           |
| ARPE           | Human retinal pigment epithelial cell line |           |
| LensEpi        | Human lens epithelial cell line            |           |
| NHDF           | Normal Human Dermal Fibroblasts            |           |
| NHDF FBS (-)   | Normal Human Dermal Fibroblasts            |           |
| NHEK           | Keratinocyte, normal epidermal cell        |           |

(B)

| Sample name           | Description                                | Race              | Data set |
|-----------------------|--|-------------------|----------|
| GM12878               | Lymphoblast cell line                      |                   | 1        |
| H1-ES                 | Human embryonic stem cell                  |                   | 1        |
| HeLa-S3               | Cervical carcinoma cell line               |                   | 1        |
| HepG2                 | Hepatocellular carcinoma cell line         |                   | 1        |
| HUVEC                 | Human umbilical vein endothelial cell line |                   | 1        |
| K562                  | Chronic myeloid leukemia cell line         |                   | 1        |
| NHEK                  | Keratinocyte, normal epidermal cell        |                   | 1        |
| Skin                  | Human skin                                 |                   | 2        |
| trachea               | Human trachea                              | Unknown/Caucasian | 2        |
| cervix                | Human cervix                               | Unknown           | 2        |
| Brain                 | brain cerebellum                           | Asian             | 2        |
| atrium_left           | atrium_left                                | Unknown           | 2        |
| atrium_right          | atrium_right                               | Unknown           | 2        |
| bladder               | bladder                                    | Caucasian         | 2        |
| bone_marrow           | bone_marrow                                | Caucasian         | 2        |
| pancreas              | pancreas                                   | Caucasian         | 2        |
| placenta              | placenta                                   | Black             | 2        |
| small_intestine       | small_intestine                            | Unknown           | 2        |
| spleen                | spleen                                     | Unknown           | 2        |
| stomach               | stomach                                    | Unknown           | 2        |
| thymus                | thymus                                     | Caucasian         | 2        |
| uterus                | uterus                                     | Unknown           | 2        |
| fetal_brain           | fetal_brain                                | Unknown           | 2        |
| fetal_colon           | fetal_colon                                | Unknown           | 2        |
| fetal_liver           | fetal_liver                                | Unknown           | 2        |
| fetal_small_intestine | fetal_small_intestine                      | Unknown           | 2        |
| fetal_stomach         | fetal_stomach                              | Unknown           | 2        |
| IHEC6A-1_HPAEC        | IHEC6A-1_HPAEC                             |                   | 3        |
| IHEC6A-2_HDMEC        | IHEC6A-2_HDMEC                             |                   | 3        |
| IHEC6A-3_HIAEC        | IHEC6A-3_HIAEC                             |                   | 3        |
| IHEC6A-4_HCAEC        | IHEC6A-4_HCAEC                             |                   | 3        |
| IHEC6A-5_HAoEC        | IHEC6A-5_HAoEC                             |                   | 3        |
| IHEC6A-6_HUVEC        | IHEC6A-6_HUVEC                             |                   | 3        |
| IHEC6A-7_HUVEC        | IHEC6A-7_HUVEC                             |                   | 3        |
| IHEC6A-7v2_HUVEC      | IHEC6A-7v2_HUVEC                           |                   | 3        |
| IHEC6A-8_HUVEC        | IHEC6A-8_HUVEC                             |                   | 3        |
| IHEC6A-8v2_HUVEC      | IHEC6A-8v2_HUVEC                           |                   | 3        |
| IHEC7A-1_HMVEC-C      | IHEC7A-1_HMVEC-C                           |                   | 3        |
| IHEC7A-2_LMVC-Ad      | IHEC7A-2_LMVC-Ad                           |                   | 3        |
| IHEC7A-3_UtMVEC       | IHEC7A-3_UtMVEC                            |                   | 3        |
| IHEC7A-4_HMVEC-L      | IHEC7A-4_HMVEC-L                           |                   | 3        |
| IHEC7A-5_HSVEC        | IHEC7A-5_HSVEC                             |                   | 3        |

1. mRNA-seq\_ENCODE\_Duke university
2. Aburatani-lab body map project
3. The international epigenome consortium



### 3B7. 逆転写、定量リアルタイムPCR

1  $\mu\text{g}$  のトータル RNA を鋳型とし、SuperScript III Reverse Transcriptase (Invitrogen) とランダムヘキサマーを用いて、50°C にて 50 分間、逆転写を行った。反応後の cDNA は純水 79  $\mu\text{l}$  を加えて 100  $\mu\text{l}$  として、cDNA の状態で -20°C にて保存した。

上記 cDNA 溶液を 5 倍希釈し 1  $\mu\text{l}$  を定量 PCR に用いた。定量 PCR は、SYBR Green I 法にてリアルタイム PCR 装置 (CFX96 Real Time System、BioRad) にて行った。

使用したプライマーは以下の通りである。

ACTB forward: 5'-ACTGGCATGGCCTTCCGTGTTCTTA-3'

ACTB reverse: 5'-TCAGTGTAGCCCAAGATGCCTTC-3'

KRT12 forward: 5'-AGCAGAATCGGAAGGACGCTGA-3'

KRT12 reverse: 5'-ACCTCGCTCTTGCTGGACTGAA-3'

KRT3 forward: 5'-ACGTGACTACCAGGAGCTGATG-3'

KRT3 reverse: 5'-ATGCTGACAGCACTCGGACACT-3'

PAX6 forward: 5'-TGGGAAATCCGAGACAGATTACTG-3'

PAX6 reverse: 5'-TCTGCCCCGTTCAACATCCTTAG-3'

### 3B8. 免疫染色

ヒトドナー強角膜片を OCT コンパウンドにて包埋した後、液体窒素にて直ちに凍結し、-80°C にて保存した。クリオスタット (Leica) にて厚さ 8  $\mu\text{m}$  に薄切し、10%ホルムアルデヒドにて固定後、2%スキムミルクにてブロッキングを

行った。抗 KRT3 抗体 (AE-5; Progen Biotechnik GMBH) を 1000 倍希釈、抗 KRT12 抗体 (N-16; Santa cruz) を 500 倍希釈にて 4°C で 1 晩インキュベートした。PBS にて洗浄後、Alexa Fluor 488 抗マウス二次抗体 (#A-11001、Molecular Probes) を用いて 1 時間室温で静置後、洗浄して、DAPI 入り封入剤 (Vectashield) で封入した。蛍光顕微鏡 (BZ-8100、キーエンス) を用いて撮影した。

### 3C. 結果

#### 3C1. ドナー強角膜片より得られたRNAの量と質

ヒトドナー強角膜片より抽出した RNA サンプルの濃度 (ng/μl) と質 (RNA integrity number、RIN) を表 4 に示す。トータル RNA は全てのサンプルで 12 μl の純水に溶出した。量は全てのサンプルで 100 ng 以上採取され、質は全てのサンプルで、RIN 値 7 以上の基準を満たし、網羅的な遺伝子発現解析が可能であった (表 4)。

表 4 : 網羅的発現解析に用いたヒトドナー角膜上皮細胞の RNA 濃度と質

| Assay name   | RNA concentration<br>(ng / $\mu$ l) | RNA integrity<br>number (RIN) |
|--------------|-------------------------------------|-------------------------------|
| RNA-seq_1    | 30                                  | 8.6                           |
| RNA-seq_2    | 111                                 | 10                            |
| Exon array_1 | 26                                  | 8.6                           |
| Exon array_2 | 38                                  | 8.6                           |

### 3C2. Exon arrayを用いた、角膜上皮特異的なマーカー遺伝子の探索

Exon array を用いて、ヒトドナー角膜上皮とヒト正常組織・培養細胞 81 サンプル（表 2）の発現を比較し、角膜上皮特異的なマーカー遺伝子を探索した。①角膜上皮の全てのサンプルで発現値（Raw signal）が皮膚の 10 倍以上、かつ②角膜上皮の全てのサンプルで発現値がヒト正常組織・培養細胞 81 サンプルの中央値の 10 倍以上、を満たす遺伝子が 52 遺伝子同定された（表 5、図 8）。

この中から、角膜上皮でのみ発現している特異性の高い遺伝子を選び出すため、52 遺伝子について、ヒト正常組織・培養細胞 81 サンプルのうち角膜上皮の発現値（Raw signal）の 0.1 倍以上の組織・細胞の数を算出した。52 遺伝子を、角膜上皮の発現値（Raw signal）の 0.1 倍以上の組織・細胞の数が少ない順にソートして示す（表 5、図 8）。

ヒト正常組織・培養細胞 81 サンプルのうち角膜上皮の発現値（Raw signal）の 0.1 倍以上の組織・細胞の数が 0 個であるのは KRT12 のみであった（表 5、図 8、図 9A）。また 1 個である遺伝子は、PAX6、KRT3 だった（表 5、図 9）。KRT3 の発現は舌にのみに認められた（図 9B）。PAX6 は脳で発現を認めた（図 9C）。KRT24 は、食道、子宮頸部、胎盤で発現を認めた（図 9D）。KRT12 は角膜上皮で特異的に発現していることが確認された（図 9A）。タンパク質レベルでも、免疫染色により KRT12、KRT3 の角膜上皮における発現が確認された（図 10）。

表 5 : Exon array を用いた角膜上皮特異的遺伝子の探索

①角膜上皮の全てのサンプルで発現値 (Raw signal) が皮膚の 10 倍以上、かつ②角膜上皮の全てのサンプルで発現値がヒト正常組織 81 サンプルの中央値の 10 倍以上を満たす 52 遺伝子を、# of tissues > 0.1 × CEpi in 81 tissues & cells が小さい順に示す。

# of tissues > 0.1 × CEpi in 81 tissues & cells: ヒト正常組織 81 サンプルのうち、角膜上皮の発現値 (Raw signal) の 0.1 倍以上のサンプル数。FC\_CEpi/ Skin: 角膜上皮/ 皮膚の発現比。FC\_CEpi/ median (81 tissues & cells): 角膜上皮/ ヒト正常組織 81 サンプルの中央値の発現比。

| # of tissues > 0.1 X CEpi in 81 tissues & cells | FC_CEpi/ Skin | FC_CEpi/ median (81 tissues & cells) | Gene name | Gene annotation                                 |
|---|---------------|--------------------------------------|-----------|---|
| 0   | 354           | 503                                  | KRT12     | keratin 12                                      |
| 1   | 212           | 163                                  | PAX6      | paired box gene 6 isoform a                     |
| 1   | 23            | 32                                   | KRT3      | keratin 3                                       |
| 2   | 63            | 65                                   | KRT24     | keratin 24                                      |
| 3   | 27            | 79                                   | SOX15     | SRY-box 15                                      |
| 3   | 29            | 28                                   | ARL4D     | ADP-ribosylation factor-like 4D                 |
| 4   | 133           | 136                                  | UPK1B     | uroplakin 1B                                    |
| 4   | 40            | 87                                   | FAM83A    | hypothetical protein LOC84985 isoform a         |
| 4   | 52            | 36                                   | OSAP      | ovary-specific acidic protein                   |
| 5   | 64            | 121                                  | ALDH3A1   | aldehyde dehydrogenase 3 family, member A1      |
| 5   | 19            | 37                                   | MAB21L1   | mab-21-like protein 1                           |
| 6   | 89            | 89                                   | TMPRSS11D | transmembrane protease, serine 11D              |
| 6   | 52            | 37                                   | C6orf205  | mucin 21  |
| 8   | 21            | 17                                   | C8orf47   | hypothetical protein LOC203111                  |
| 8   | 78            | 52                                   | CRTAC1    | cartilage acidic protein 1                      |
| 9   | 20            | 17                                   | ULBP2     | UL16 binding protein 2                          |
| 10  | 52            | 129                                  | TMPRSS4   | transmembrane protease, serine 4 isoform 3      |
| 13  | 30            | 24                                   | HRASL2    | HRAS-like suppressor 2                          |
| 14  | 16            | 15                                   | RAET1G    | retinoic acid early transcript 1G               |
| 15  | 11            | 18                                   | EVPL      | envoplakin                                      |
| 16  | 12            | 58                                   | GPR109B   | G protein-coupled receptor 109B                 |
| 16  | 11            | 15                                   | PHLDA2    | pleckstrin homology-like domain family A member |
| 16  | 12            | 14                                   | ADH7      | class IV alcohol dehydrogenase 7 mu or sigma    |
| 17  | 64            | 40                                   | TRIM36    | tripartite motif-containing 36 isoform 2        |
| 17  | 23            | 27                                   | XDH       | xanthine dehydrogenase                          |
| 17  | 13            | 13                                   | HORMAD1   | HORMA domain containing 1                       |
| 18  | 11            | 16                                   | GDPD2     | osteoblast differentiation promoting factor     |
| 18  | 24            | 28                                   | RASSF6    | Ras association (RalGDS/AF-6) domain family 6   |
| 20  | 23            | 25                                   | ELF3      | E74-like factor 3 (ets domain transcription     |
| 21  | 22            | 14                                   | CSRP2     | cysteine and glycine-rich protein 2             |
| 22  | 104           | 20                                   | NQO1      | NAD(P)H menadione oxidoreductase 1,             |
| 22  | 11            | 11                                   | SSTR5     | somatostatin receptor 5                         |
| 25  | 10            | 14                                   | DUSP7     | dual specificity phosphatase 7                  |
| 25  | 16            | 12                                   | APOBEC3A  | phorbolin 1                                     |
| 26  | 14            | 13                                   | IER3      | immediate early response 3                      |
| 27  | 14            | 11                                   | WNT7A     | wingless-type MMTV integration site family,     |
| 28  | 10            | 15                                   | NR1D1     | nuclear receptor subfamily 1, group D, member 1 |
| 28  | 23            | 13                                   | TGFB1     | transforming growth factor, beta-induced, 68kDa |
| 29  | 28            | 13                                   | SCARA3    | scavenger receptor class A, member 3 isoform 1  |
| 30  | 23            | 13                                   | CPAMDB    | C3 and PZP-like, alpha-2-macroglobulin domain   |
| 33  | 20            | 12                                   | ULBP3     | UL16 binding protein 3                          |
| 33  | 13            | 11                                   | FOSL1     | FOS-like antigen 1                              |
| 34  | 15            | 12                                   | WNT7B     | wingless-type MMTV integration site family,     |
| 34  | 13            | 12                                   | GPR110    | G-protein coupled receptor 110 isoform 2        |
| 35  | 21            | 11                                   | CACNG4    | voltage-dependent calcium channel gamma-4       |
| 35  | 40            | 14                                   | TSPAN1    | tetraspan 1                                     |
| 36  | 17            | 11                                   | BCL3      | B-cell CLL/lymphoma 3                           |
| 37  | 12            | 10                                   | MALL      | mal, T-cell differentiation protein-like        |
| 38  | 12            | 11                                   | NUAK2     | NUAK family, SNF1-like kinase, 2                |
| 38  | 11            | 12                                   | SCNN1A    | sodium channel, nonvoltage-gated 1 alpha        |
| 38  | 16            | 12                                   | AGR2      | anterior gradient 2 homolog                     |
| 40  | 19            | 10                                   | CYP2S1    | cytochrome P450, family 2, subfamily S,         |

# 角膜

# ヒト正常組織81サンプル

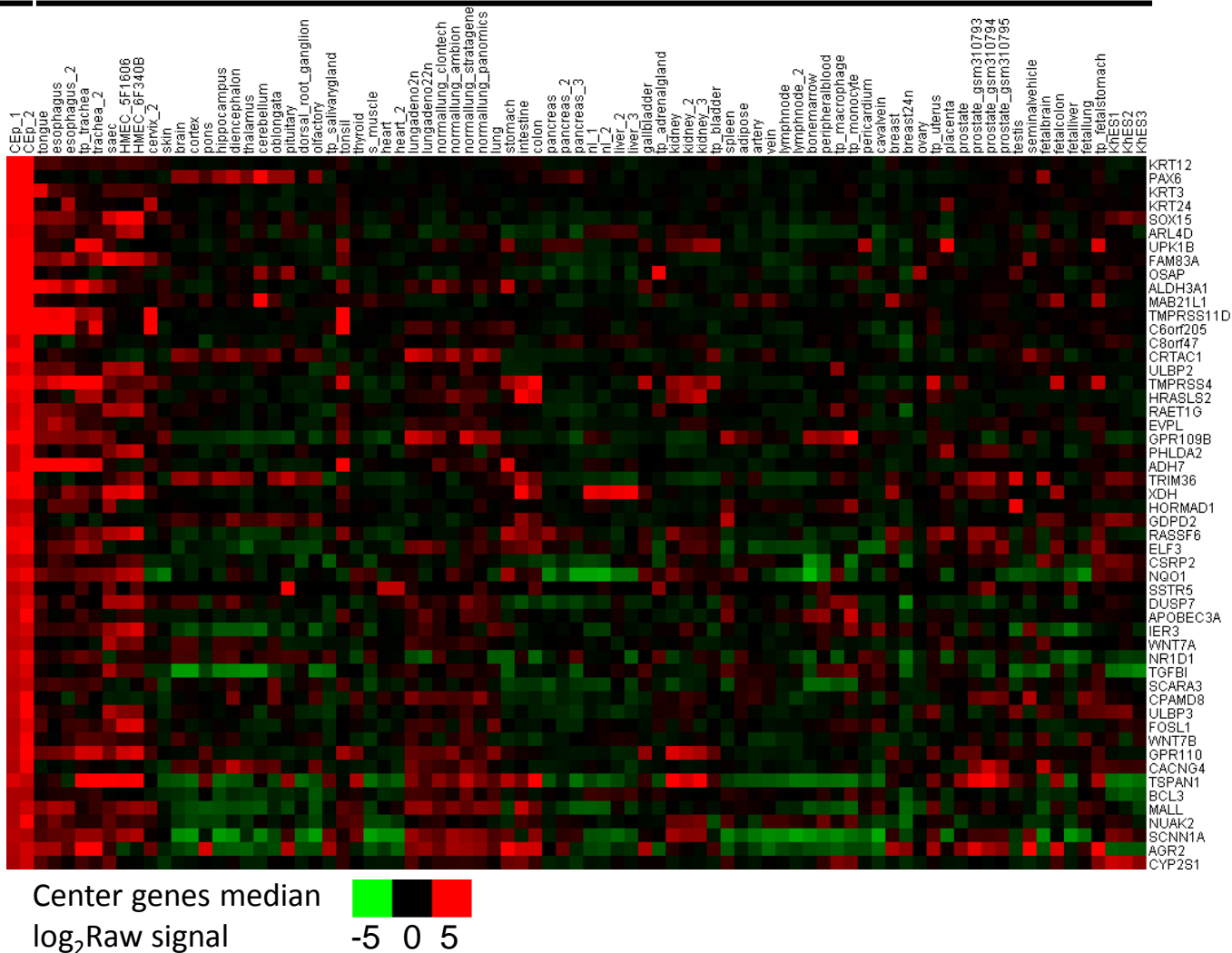


図 8 : 網羅的発現解析 (Exon array) を用いた 52 の角膜上皮特異的遺伝子  
 表 6 と同じ 52 遺伝子をヒートマップで示す。左 2 列がヒトドナー角膜上皮細胞、残り 81 列が比較対象のヒト正常組織・培養細胞 81 サンプルを示す。

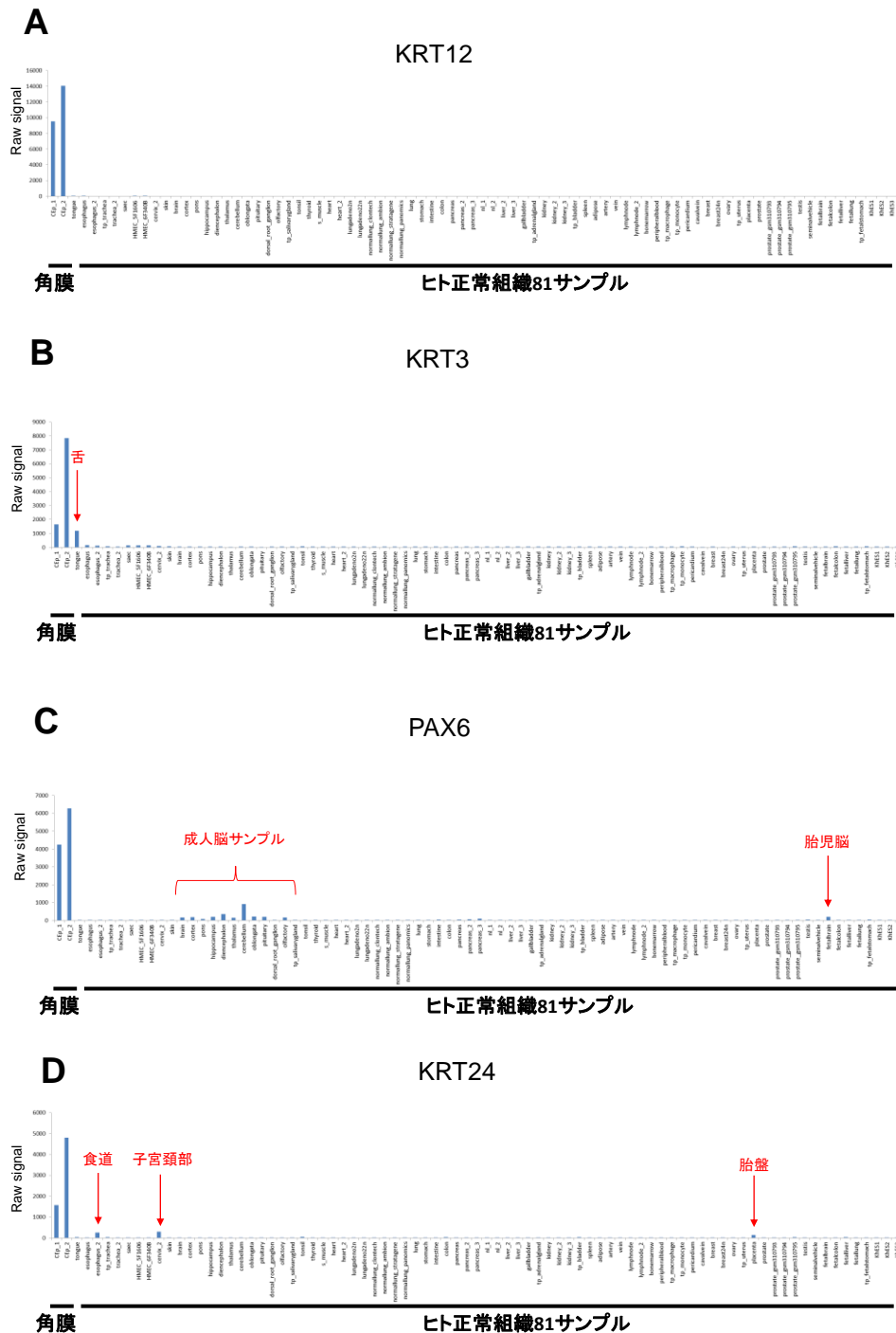
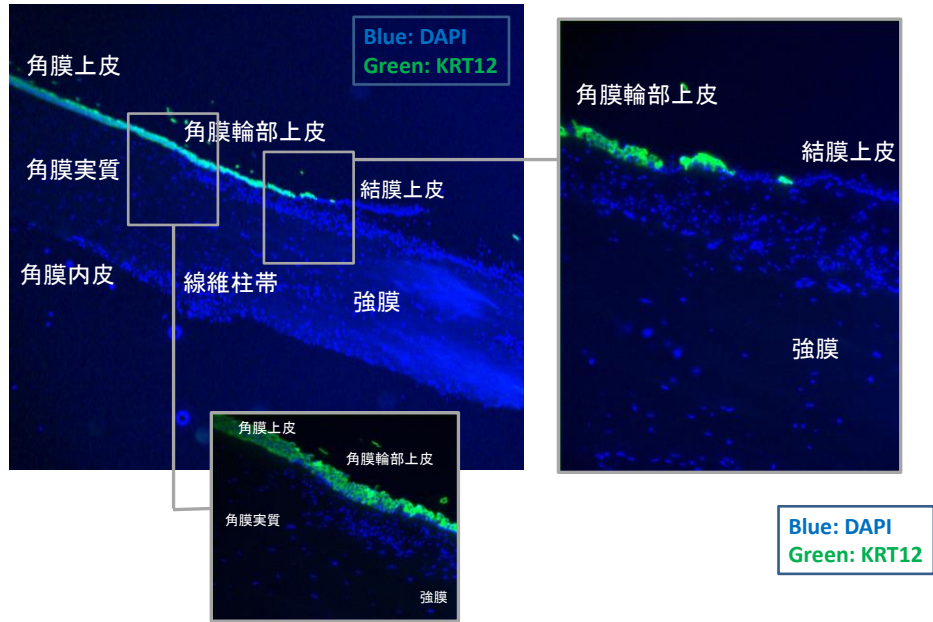


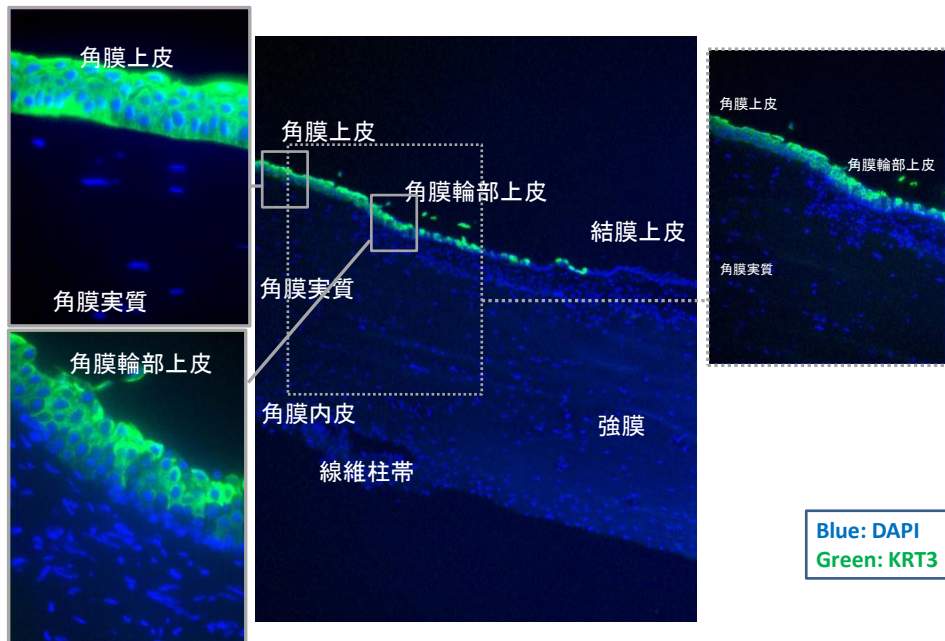
図 9: KRT12、KRT3、PAX6、KRT24 の角膜上皮とヒト正常組織 81 サンプルの発現 (Exon array)

角膜上皮 (CEpi1、CEpi2) は KRT12、KRT3、PAX6、KRT24 を特異的に発現していた。ヒト正常組織 81 サンプルでは、KRT12 を発現しているサンプルは無かった (図 9A)。ヒト正常組織 81 サンプルでは、舌において、KRT3 の発現を認めた (図 9B)。PAX6 は成人脳で発現を認めた (図 7C)。KRT24 は食道、子宮頸部、胎盤で発現を認めた (図 9D)。

A



B





C

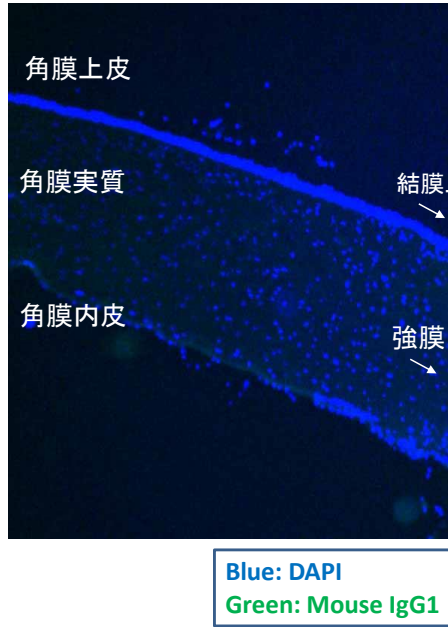


図 10: ヒトドナー強角膜片における KRT12、KRT3 の免疫染色像

青色：核染色（DAPI）。緑色：KRT12（図 10A）、KRT3（図 10B）、陰性コントロール（図 10C）。

### 3C3. mRNA-seqを用いた初代培養角膜輪部上皮細胞、角膜上皮細胞株の遺伝子発現解析

*in vitro* での角膜上皮の特異性を制御する転写因子 (KRT12 を制御する転写因子) の研究、ノックダウン実験や、細胞数を要するゲノムワイドなエピゲノム解析のためには、角膜上皮の表現型 (たとえば、KRT12 を発現する) を保持する培養細胞が必要となる。しかし、広く用いられている SV40 large T 抗原不死化角膜上皮細胞の細胞株<sup>51,54</sup> の KRT3、KRT12 の発現は明らかでない。そのため、初代培養角膜輪部上皮細胞と角膜上皮細胞株の角膜上皮特異的遺伝子の発現を調べた。

まず、初代培養角膜輪部上皮細胞の発現を調べた。培養法としては、以前我々が報告した無血清・無フィーダー培養を用いた<sup>31,32</sup>。同方法で作成した培養シートは、生体内と同様に4~5層に重層化していた (図 11)。mRNA-seq で遺伝子発現を調べたところ、初代培養細胞において発現 (FPKM>1) が確認できたのは、52 個の角膜上皮特異的遺伝子のうち 47 個であり (図 13)、KRT3、KRT12、PAX6 のいずれも発現が確認された (図 12)。一方、細胞株において発現 (FPKM>1) が確認できたのは、52 個の角膜上皮特異的遺伝子のうち 30 個に留まり (図 13)、KRT3、KRT12、PAX6 は発現していなかった (図 12)。細胞株を初代培養細胞と同じ無血清・無フィーダーの培養条件<sup>31,32</sup> で培養しても、KRT3、KRT12、PAX6 の発現は認められなかった (図 14)。

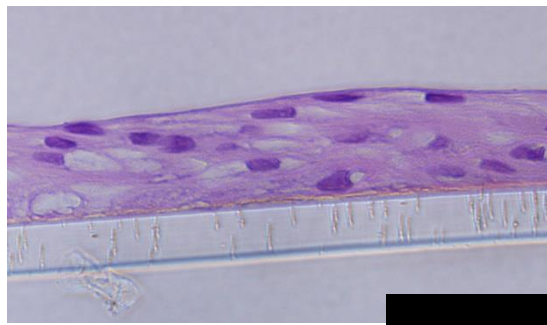


図 11: 初代培養角膜輪部上皮細胞シート  
HE 染色。Bar = 50  $\mu\text{m}$ 。

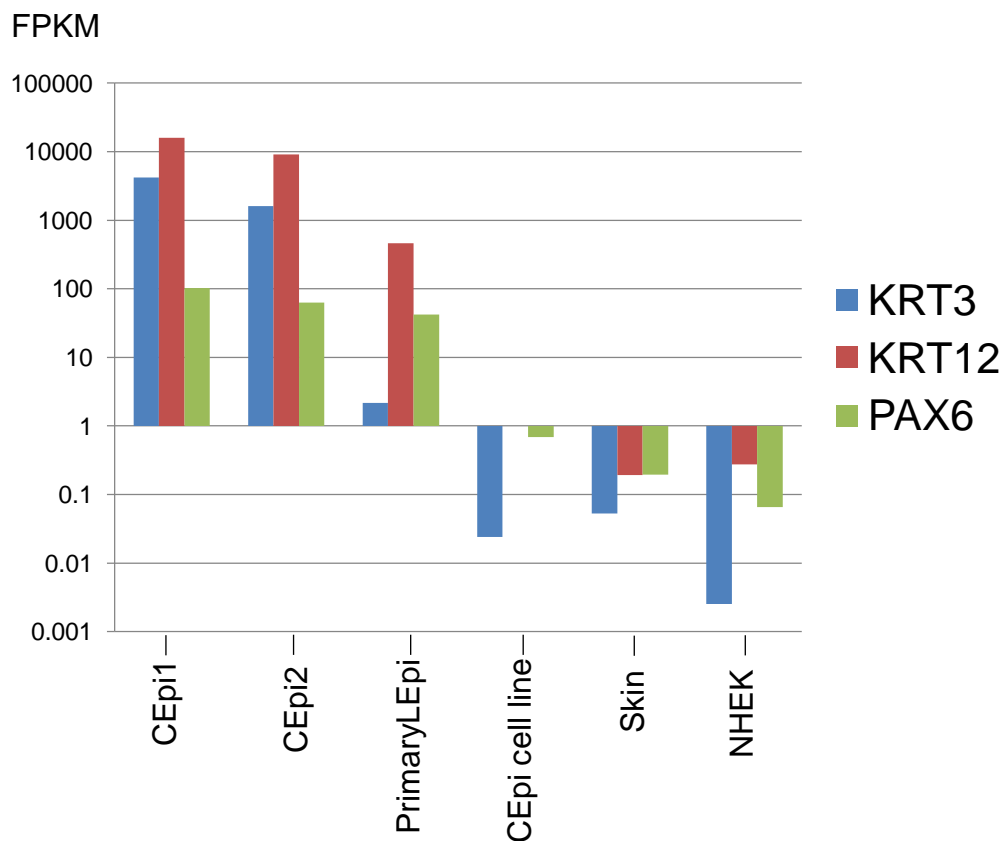


図 12: 初代培養角膜輪部上皮細胞、角膜上皮細胞株における KRT3、KRT12、PAX6 の遺伝子発現 (mRNA-seq)

発現値は FPKM で示した。CEpi: ヒトドナー角膜上皮細胞、PrimaryLEpi: 初代培養角膜輪部上皮細胞、CEpi cell line: 角膜上皮細胞株、Skin: ヒト皮膚、NHEK: 正常ヒト表皮角化細胞。

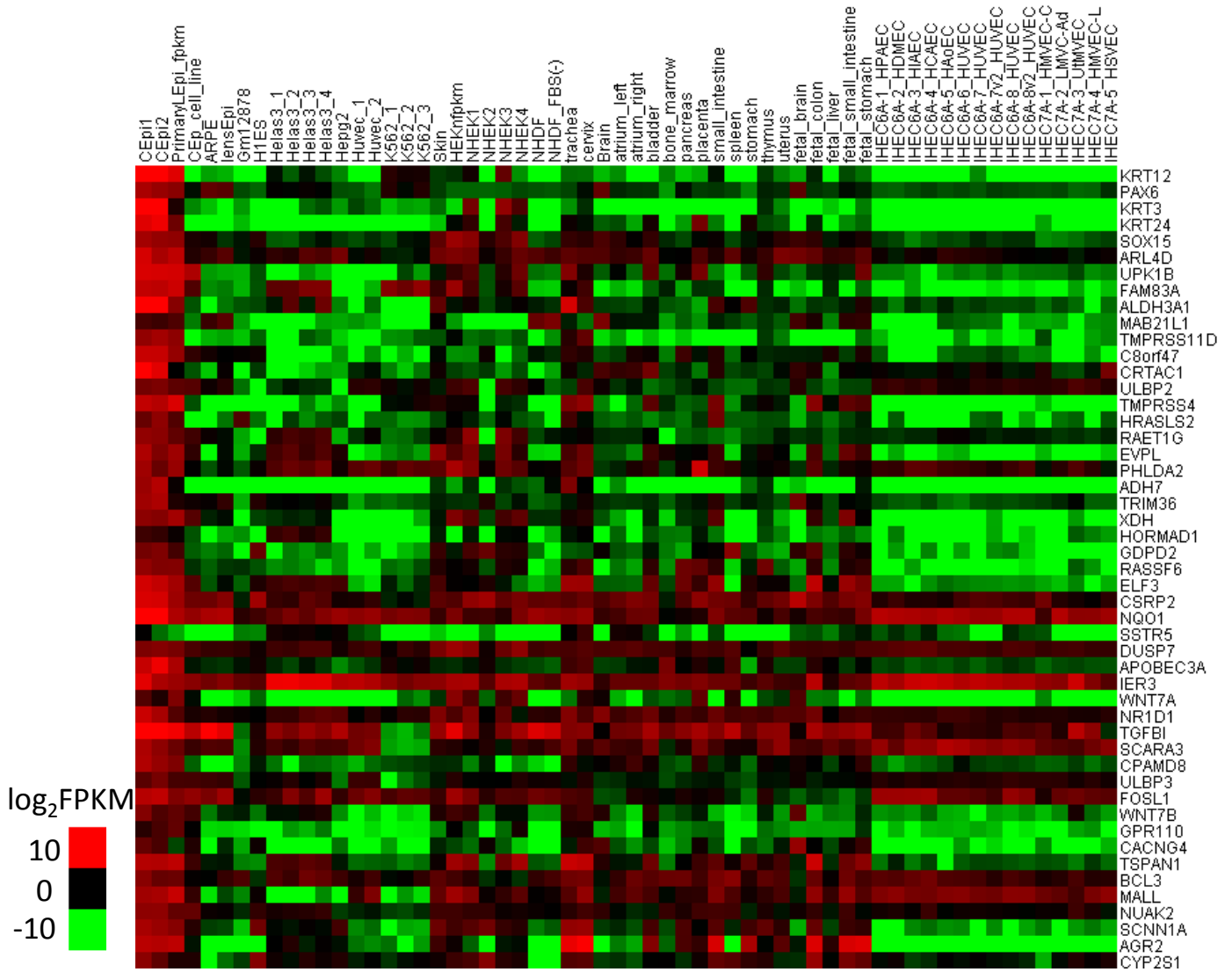
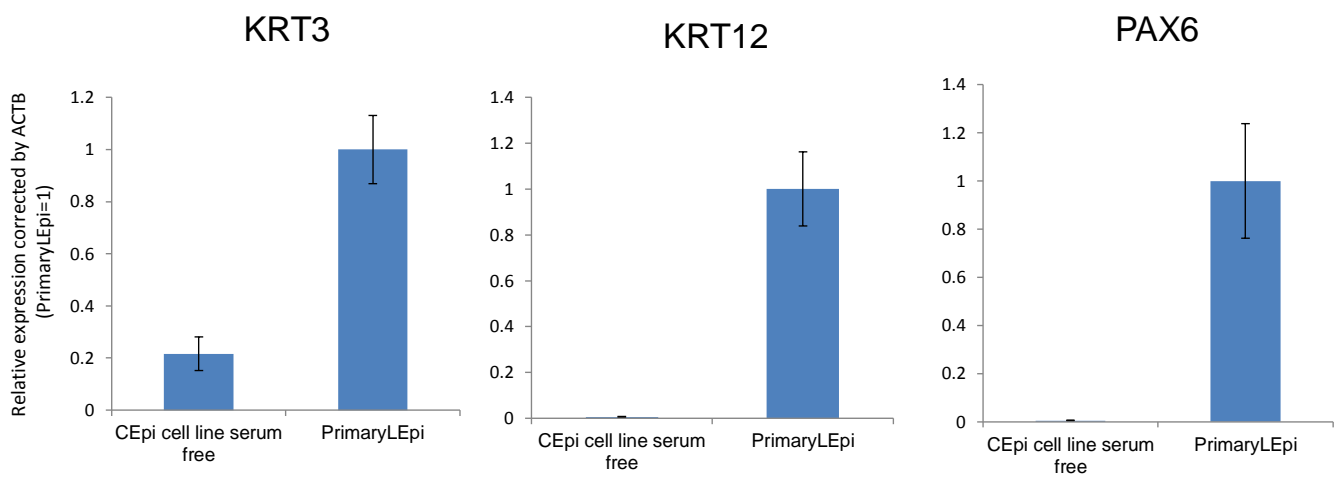


図 13: mRNA-seq を用いたヒトドナー角膜上皮細胞、初代培養角膜輪部上皮細胞、角膜上皮細胞株における 52 個の角膜上皮特異的遺伝子の発現

発現値は FPKM で示した。左より、CEpi1、CEpi2: ヒトドナー角膜上皮細胞、PrimaryLEpi\_fpkm: 初代培養角膜輪部上皮細胞、CEpi cell\_line: 角膜上皮細胞株。



**図 14: 角膜上皮細胞株の無血清無血清無フィーダー培養における KRT3、KRT12、PAX6 の遺伝子発現**

qRT-PCR による遺伝子発現 (ACTB で補正、初代培養角膜輪部上皮細胞 = 1) を示す。CEpi cell line serum free: insert culture 上で無血清・無フィーダー培養した角膜上皮細胞株、PrimaryLEpi: 初代培養角膜輪部上皮細胞。

### 3C4. mRNA-seq、Exon arrayを用いた皮膚と比較して角膜上皮で発現が高い転写因子の探索

前項までにより、角膜上皮の特異性が明らかになった。KRT12が最も特異性のマーカーであることが確認され、KRT3、PAX6が次いで特異性の高いマーカーと考えられた。また、これらには落ちるものの、計52個の角膜上皮で特異的に発現上昇している遺伝子を同定した。これら52個のうち、mRNA-seqでアノテーションできたものは49個あり、49個中、SSTR5をのぞいた48個がmRNA-seqでもヒトドナー角膜上皮細胞における発現(FPKM > 1)を確認できた(図13)。

次に、このような角膜上皮の特異性を規定している転写因子を明らかにするため、まず、皮膚と比較して角膜で発現が高い転写因子を同定した。

まず、mRNA-seqを用いて、角膜上皮の発現値(FPKM)をヒト皮膚組織の発現値と比較し、皮膚と比較して角膜上皮で発現が高い転写因子を抽出した。本研究における転写因子は、Heinäniemi Mらが定義した計2,754個の転写制御に関わると考えられている遺伝子(Transcription factor、Co-regulator、Chromatin modifier、mRNA transcript synthesis/processing、domain-based evidence)とした<sup>55</sup>。これらのうち、①角膜上皮の全てのサンプルで発現値(FPKM)が皮膚の2倍以上、かつ②角膜上皮の全てのサンプルでFPKMが1以上、を満たす転写因子が88個同定された(図15)。

次に、Exon arrayを用いて、①角膜上皮の全てのサンプルで発現値(Raw signal)が皮膚の2倍以上、②角膜上皮の全てのサンプルでRaw signalが100以上、を満たす転写因子を755個同定した(図15)。

以上のmRNA-seqとExon arrayの両方で発現が高いことが確認できた転写因

子は 59 個あった。59 転写因子を図 15B にリストで示す。



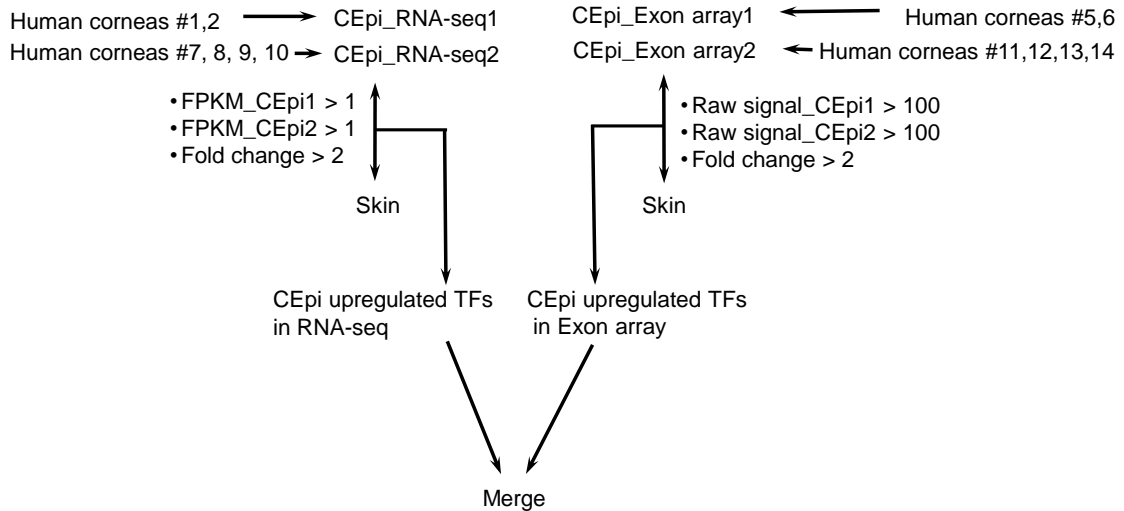
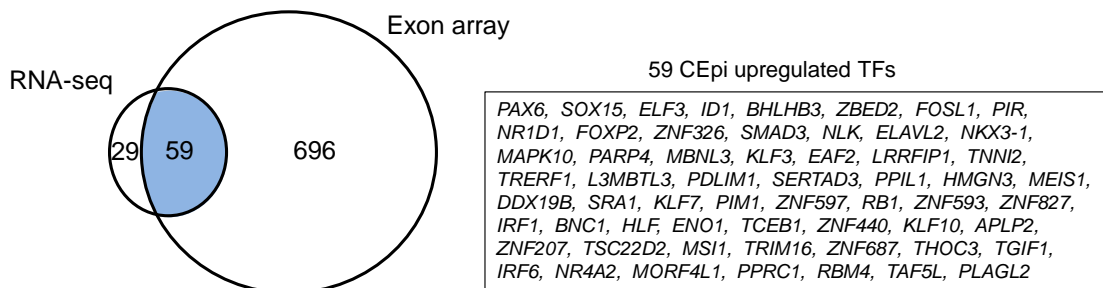
**A****B**

図 15: 皮膚に比べて角膜上皮で発現の高い転写因子の探索

(A) 実験概要。RNA-seq と Exon array で共通して発現が上昇している転写因子を選択した。(B) RNA-seq と Exon array により同定された角膜上皮特異的な転写因子の数とその内訳。CEpi: Human donor corneal epithelium, Skin: human skin, TFs: transcription factors (Heinäniemi M et al. Nature Methods 2013)。

## IV. 角膜上皮細胞のクロマチン構造解析

### 4A. 背景

ヌクレオソーム・フリー領域をゲノムワイドに同定する Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) -sequencing とモチーフ解析について説明する。

#### 4A1. ヌクレオソーム・フリー領域

遺伝子発現の物理的制御としては、一次（ヌクレオソーム） および高次よりなるクロマチン構造が重要である。クロマチン構造は正常な細胞分化や発生の過程でダイナミックに変化する（クロマチンリモデリング）。代表的な細胞種におけるヒストン修飾状態、クロマチン構造のプロファイルは、2000年代に発足した ENCODE (The Encyclopedia of DNA Elements) プロジェクトで公開されており、多くの生物の転写制御の解明に広く貢献してきた<sup>53,56</sup>。各種臓器の培養細胞株を用いて行われた FAIRE および DNase-seq のデータの統合解析においても、細胞株特有のヌクレオソーム・フリー領域は、エンハンサー領域と重なることが示されている<sup>57,58</sup>。これら近年の研究結果は、分化の過程において、特にエンハンサー領域の転写因子の結合とそれにとまなうヒストン修飾やヌクレオソーム・フリー領域の変化が重要であることを示唆している<sup>59-61</sup>。遺伝子発現状態は転写開始点が存在するプロモーター領域のクロマチン状態とよく相関するが、細胞系譜に特有の遺伝子発現調節においては、エンハンサーの

クロマチン状態も深く関与していることが知られている<sup>62</sup>。

#### 4A2. FAIRE-seq法

Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) 法は、ホルムアルデヒド固定した細胞核画分に対してフェノール・クロロホルム精製を行い、DNA・タンパク質架橋が起こりにくいヌクレオソーム・フリー領域のDNAを水層に分離することで、これを回収、濃縮するという手法である。これに次世代シーケンサーを組み合わせ、ゲノムワイドにヌクレオソーム・フリー領域を同定するのがFAIRE-seq法である<sup>63,64</sup>。

#### 4A3. モチーフ解析

転写因子は、固有のDNA結合ドメインを介した特定のDNA配列（シスエレメント）に対する親和性が強く、二本鎖DNAとの結合と解離を繰り返しながら、コアヒストンのないヌクレオソーム・フリー領域を形成している<sup>65</sup>。モチーフ解析は、このヌクレオソーム・フリー領域の配列解析を行い、その領域に結合する転写因子を予測する手法である<sup>57</sup>。

## 4B. 方法

### 4B1. 概要

ドナー角膜上皮細胞 (CEpi) の FAIRE-seq を行い、ヌクレオソーム・フリー領域を同定した。FAIRE-seq で用いたドナー角膜のプロファイルを表 1 に示す。FAIRE\_1 は 2 個人の両眼、計 4 眼の角膜を用いて FAIRE-seq を行った。FAIRE\_2 は、1 個人の 1 眼を用いて FAIRE-seq を行った。解析には、FAIRE\_1 を用い、FAIRE\_2 にて、再現性を確認した。同時にヒト網膜色素上皮細胞株 (ARPE19)、ヒト水晶体上皮細胞株 (lensEpi) の 2 サンプルについて FAIRE-seq を行い、また ENCODE プロジェクト<sup>53</sup> の 7 サンプルの DNase-seq (リンパ芽球細胞株 (GM12878)、ヒト胚性幹細胞 (H1-ES)、子宮頸癌細胞株 (HeLa-S3)、肝細胞癌細胞株 (HepG2)、ヒト臍帯静脈内皮細胞 (HUVEC)、慢性骨髄性白血病細胞株 (K562)、正常ヒト表皮角化細胞 (NHEK) (Washington 大学) を用い、計 9 サンプルを角膜上皮の比較対象として用いた (表 6)。

### 4B2. FAIRE

FAIRE は培養細胞のプロトコル<sup>66</sup>に準拠して行った。以下に具体的に述べる。

FAIRE に用いたドナー角膜上皮細胞 (CEpi) は、研究用ヒト強角膜片を実体顕微鏡 (SZX12、Olympus) 下で機械的に剥離し、PBS 中に採取した (図 5)。得られた細胞懸濁液に、最終濃度が 1% になるようにホルムアルデヒドを加え、室温で 10 分間の固定 (クロスリンク) を行った。次に、最終濃度が

0.125 mM になるようにグリシン溶液を添加し反応を停止させ、4°C、2,000 × g で遠心してペレットを得た。ペレットに 600 μl の核溶解バッファー（10 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM EDTA, 1% SDS）を添加してクロマチンを抽出した。これを、超音波細胞破碎装置（UD-201、トミー精工）を用いて氷上で破碎し、平均長が 200~500 塩基となるようにクロマチン DNA の断片化を行った。断片化したサンプルを 4°C、15000 × g で 15 分間遠心し上清をクロマチン溶液とした。クロマチン溶液に等量の PCI 溶液（フェノール:クロロホルム:イソアミルアルコール=25:24:1）を混合し、25°C、15000 × g で 25 分間遠心の遠心分離を行い、水層を回収した。同様に等量の CIA 溶液（クロロホルム:イソアミルアルコール=24:1）を混合し、遠心分離した後、水層を DNA 溶液として回収した。DNA 溶液に 1/10 量の 3M 酢酸ナトリウム（pH 5.2）と 2.5 倍量の 100%エタノールを混合し、4°C、15000 × g で 30 分間遠心し、DNA を沈殿させた。沈殿を 70%エタノールでリンスした後、1 μg の RNaseA を添加し、37°C で 30 分間静置した。さらに 1 μg のプロナーゼを添加し、56°C で 1 時間静置しタンパク質を除去した。このサンプルを QIAquick PCR purification Kit（Qiagen）を用いて精製した。

インプットサンプルとして、上記のクロマチン溶液を用い、同様に RNaseA およびプロナーゼ処理した。65°C で一晩脱クロスリンク反応を行った。PCI 溶液および CIA 溶液を用いて水層を回収し、エタノール沈殿後、QIAquick PCR purification Kit を用いて精製した。

精製したサンプルは、Qubit dsDNA HS Assay Kit と Qubit 2.0 Fluorometer

(Life Technologies) を用いて濃度を測定した。

比較対象実験として、ヒト網膜色素上皮細胞株 (ARPE19)、ヒト水晶体上皮細胞株 (lensEpi) についても同様に FAIRE サンプルを調整した。

#### 4B3. FAIRE サンプルの濃縮確認

FAIRE 法で得られたサンプルは、シーケンスの前に定量 PCR を用いて、その濃縮率を比較・評価した。一般的に、この濃縮率が 10 倍以上では、良好なシーケンスデータを得られる可能性が高く、逆に 5 倍以下では、ピークの同定の困難なバックグラウンドの高いデータとなる。

具体的には、得られた FAIRE サンプルの濃縮の確認 をリアルタイム qPCR にて行った。ポジティブコントロール領域は GAPDH のプロモーター領域、ネガティブコントロール領域は HBB の遺伝子本体領域とした。SYBR Green I 法を用いてリアルタイム PCR 装置にて FAIRE サンプル、INPUT サンプルそれぞれについて GAPDH プロモーター領域、HBB 遺伝子本体領域の Starting quantity (SQ) を求めた。結合領域 (陽性コントロール) のプライマー、および内部標準領域 (陰性コントロール) のプライマーは下記の通りである。

GAPDH forward: 5'-CACGTAGCTCAGGCCTCAAGA-3'

GAPDH reverse: 5'-GGCTGCGGGCTCAATTTAT-3'

HBB forward: 5'-GGGCTGAGGGTTTGAAGTCC-3'

HBB reverse: 5'-CCACAGGGTGAGGTCTAAGTG-3'

次に以下の式にて濃縮率を算出した。

$$\text{濃縮率} = \frac{\frac{\text{SQ (GAPDH}_{\text{FAIRE}})}{\text{SQ (GAPDH}_{\text{INPUT}})}}{\frac{\text{SQ (HBB}_{\text{FAIRE}})}{\text{SQ (HBB}_{\text{INPUT}})}}$$

#### 4B4. 次世代シーケンサーによるFAIREサンプルの解析

5倍以上の良好な濃縮を得られた FAIRE サンプルを、次世代シーケンサー Illumina GAIIx (Illumina) で解析した。ライブラリの作成や Illumina GAIIx の解析手順は既報に準じて、以下のように行った<sup>67</sup>。

#### 4B5. シングルエンドライブラリー作成

精製された FAIRE サンプルの DNA 末端を平滑化した。QIAquick PCR purification Kit を用いて DNA を精製した後、FAIRE サンプルの 3'末端に terminal transferase で A を付加した。MinElute PCR purification Kit (Qiagen) を用いて精製後、2%アガロースゲルを用いて 100 V で 70 分間泳動を行った。サイズが 275~325 bp、375~425 bp、475~525 bp に相当する部分を切り出し、QIAquick Gel Extraction Kit (Qiagen) を用いて精製した。この DNA サンプルに対して、18 サイクル (98°C 10 秒間、65°C 30 秒間、72°C 30 秒間を 1 サイクルとする) の PCR 反応を行った。PCR 産物を MinElute PCR purification Kit を用いて精製して、シーケンス用ライブラリとした。

#### 4B6. ライブラリの確認

作成したライブラリ 1 µl を使用して、バイオアナライザによってライブラリ

の有効濃度とサイズを確認した。さらに、増幅後のライブラリについても、作成前と同様の手法を用いて定量 PCR による濃縮評価を行い、ライブラリの段階でも濃縮が得られていることを確認した。

#### **4B7. 次世代シーケンサーでの解析**

濃度および濃縮を確認したサンプルを Genome Analyzer IIx (Illumina) で解析した。リード長は 36 bp のシングルエンドで行った。

#### **4B8. ヒトリアレンスゲノムへのマッピングとピークの抽出**

Genome Analyzer IIx によって配列決定された 36 bp のシングルエンドのショートリードを、CASAVA 1.7 ソフトウェア (Illumina) によってヒトリアレンスゲノム (Hg19) にマッピングし、シーケンス配列 (タグ) の位置を決定した。これらのタグのヒトゲノム上での分布を MACS プログラム (バージョン 1.4.2) によって統計学的に検定し、P 値  $< 1e-5$  の領域を有意に濃縮した領域と判定した<sup>68</sup>。濃縮の程度を表す P 値は、IGV (Integrative Genomics Viewer, Broad institute) を用いて、ピークとして可視化した<sup>69</sup>。

#### **4B9. 転写開始点からの位置によるヌクレオソーム・フリー領域の遺伝子アノテーション**

$P < 1e-5$  を満たす各ピーク領域の遺伝子アノテーションは、最も転写開始点 (TSS: Transcription Start Site) の近い遺伝子に割り当てた (最大 1M bp 以内)



<sup>70</sup>。TSS より 2 kbp 以内にあるものプロモーター領域、それより遠方にあるものをエンハンサー領域として分類した。

#### **4B10. 公開データの利用によるDNase-seq、CTCF ChIP-seq**

7 サンプルの DNase seq は、ENCODE プロジェクトの Washington 大学の DNase-seq (GM12878、H1-ES、HeLa-S3、HepG2、HUVEC、K562、NHEK) を用いた。CTCF ChIP-seq は、ENCODE の 7 サンプル (H1-ES、GM12878、HeLa-S3、HepG2、HUVEC、K562、NHEK) を用いた。リピート配列は、UCSC サイトよりダウンロードしたものを用いた (表 6)。

データは ENCODE で公開されている配列データを、ダウンロードサイト (<http://genome.ucsc.edu/ENCODE/downloads.html>) より入手して、同じ解析方法・基準 (MACS プログラムのバージョン 1.4.2、P 値 < 1e-5 を有意) で行った。

表 6 : 本研究で用いた FAIRE-seq の材料(A)と、比較対照として用いたヒト7細胞の DNase-seq、CTCF\_ChIP-seq、Repeat sequences のデータセット(B)

(A)

| Sample name | Description                                  |
|-------------|--|
| CEpi1       | Human donor corneal epithelial cell          |
| CEpi2       | Human donor corneal epithelial cell          |
| ARPE19      | Human retinal pigmented epithelial cell line |
| LensEpi     | Human lens epithelial cell line              |

(B)

| Sample name      | Description                                | Data-set                                   |
|------------------|--|--|
| GM12878          | Lymphoblast cell line                      | DNase-seq_ENCODE_Washington university     |
| H1-ES            | Human embryonic stem cell                  | DNase-seq_ENCODE_Washington university     |
| HeLa-S3          | Cervical carcinoma cell line               | DNase-seq_ENCODE_Washington university     |
| HepG2            | Hepatocellular carcinoma cell line         | DNase-seq_ENCODE_Washington university     |
| HUVEC            | Human umbilical vein endothelial cell line | DNase-seq_ENCODE_Washington university     |
| K562             | Chronic myeloid leukemia cell line         | DNase-seq_ENCODE_Washington university     |
| NHEK             | Keratinocyte, normal epidermal cell        | DNase-seq_ENCODE_Washington university     |
| GM12878          | Lymphoblast cell line                      | CTCF_ChIP-seq_ENCODE_Washington university |
| H1-ES            | Human embryonic stem cell                  | CTCF_ChIP-seq_ENCODE_HAIB                  |
| HeLa-S3          | Cervical carcinoma cell line               | CTCF_ChIP-seq_ENCODE_Washington university |
| HepG2            | Hepatocellular carcinoma cell line         | CTCF_ChIP-seq_ENCODE_Washington university |
| HUVEC            | Human umbilical vein endothelial cell line | CTCF_ChIP-seq_ENCODE_Washington university |
| K562             | Chronic myeloid leukemia cell line         | CTCF_ChIP-seq_ENCODE_Washington university |
| NHEK             | Keratinocyte, normal epidermal cell        | CTCF_ChIP-seq_ENCODE_Washington university |
| Repeat sequences | simple sequence repeat                     | NCBI/ UCSC                                 |

#### 4B11. 角膜上皮特異的エンハンサー領域の同定

細胞の特異性を決めている領域は、Transcription start sites (TSS) 近傍のプロモーター領域よりも、遠位のエンハンサー領域である<sup>57</sup>と報告されている。FAIRE-seqにより同定されたヌクレオソーム・フリー領域から、エンハンサー領域を選択するため、以下の領域(1. TSS から 2 kb 以内<sup>57</sup>、2. NCBI に登録されているリピート配列、3. CTCF 結合領域)を除外して、エンハンサー領域を選び出した。

次に、上記のエンハンサー領域から、角膜上皮特異的な領域を、2 サンプルの FAIRE-seq (ARPE19、lensEpi) と、7 サンプルの ENCODE の DNase-seq (GM12878、H1-ES、HeLa-S3、HepG2、HUVEC、K562、NHEK)、計 9 サンプルと比較し、角膜上皮特異的エンハンサー領域を選択した。具体的には、まず、各サンプル毎に各領域の FAIRE-seq/ DNase-seq ピークの p-value を Z-Score にて算出した。次に、① 角膜上皮細胞の FAIRE-seq の Z-Score > 2.5、② 他 9 サンプルの FAIRE-seq/ DNase-seq の Z-Score < 0 であることを満たすピークを、角膜上皮特異的エンハンサー領域として選択した。

#### 4B12. モチーフ解析

角膜上皮特異的エンハンサー領域に濃縮している転写因子のモチーフ解析を行った。まず、角膜上皮特異的エンハンサー領域として選び出したピークの前 300 bp の DNA 配列を抽出した。モチーフ解析は既知のモチーフ配列の濃縮を Genomatix Overrepresented transcription factor binding sites or modules

(Genomatix, Munich, Germany)<sup>71</sup> と Transcription factor Affinity Prediction (TRAP) Web Tools<sup>72</sup> を用いて解析し、*de novo* のモチーフ解析を MOTif identification algorithm through DIrect Comparison of signal/noise distributions based on maximum entropy method (MODIC)<sup>73</sup> を用いて行った。

Genomatix を用いたモチーフ解析は、角膜上皮特異的エンハンサー領域に含まれる各転写因子ファミリーのモチーフの出現数と、同サイズのバックグラウンドに対する対象領域でのモチーフ出現数から出現比を計算し、Z-Score が高い順に Z-Score >5 を満たすモチーフを抽出した。バックグラウンドは、ヒト全染色体ゲノムとヒトプロモーターを用いた。転写因子のデータベースは MatBase<sup>71</sup> (Genomatix 社が JASPAR<sup>69</sup>、TRANSFAC<sup>70,71</sup> や既報の論文を参考にまとめた Genomatix 社オリジナルもの) を用いた。

TRAP は、角膜上皮特異的エンハンサー領域における各転写因子の結合親和性と、バックグラウンドにおける結合親和性を比較し、Benjamini-Hochberg 法による多重比較補正をした corrected P-value を求めた。バックグラウンドはヒトプロモーターを用いた。転写因子のデータベースは、JASPAR<sup>69</sup> を用いた。

MODIC は、ウィンドウサイズを 8、12、16 bp に設定し、濃縮する特徴的な転写因子の認識配列を抽出した。次に、検出したモチーフについて、STAMP<sup>74,75</sup> を用いて JASPAR<sup>76,77</sup> と TRANSFAC<sup>78,79</sup> と照合し、既知の転写因子モチーフの中で最も近い転写因子を探索した。

## 4C. 結果

### 4C1. FAIRE-seqによる角膜上皮細胞特異的なエンハンサー領域の同定

ドナー角膜上皮細胞 (CEpi) に対して、FAIRE-seq を施行した (N=2、biological replicates)。同時に、ヒト角膜上皮細胞株 (CEpi cell line)、ヒト網膜色素上皮細胞株 (ARPE19)、ヒト水晶体上皮細胞株 (lensEpi) についても FAIRE-seq を施行した。細胞株 (CEpi cell line、ARPE19、LenEpi) では、57,000~62,000 箇所のヌクレオソーム・フリー領域が同定された ( $p < 1e-5$ ) (表 7)。*In vivo* ヒトドナー角膜上皮細胞ではそれより少なく、CEpi\_1 で 35,952 箇所、CEpi\_2 で 17,326 箇所のヌクレオソーム・フリー領域が同定された。CEpi\_1 を以降の解析に用い、CEpi\_2 で再現性を確認した。CEpi\_1 で同定された 35,952 箇所のヌクレオソーム・フリー領域から、TSS より 2 kb 以内、CTCF 結合領域、リピート配列を除き、エンハンサー領域として、27,363 箇所を同定した (図 16A)。

次に、角膜上皮細胞のエンハンサー領域における、FAIRE-seq のピークの高さを Z-Score で計算した。また、ARPE19 および lensEpi の FAIRE-seq、あるいは ENCODE により取得した 7 細胞の DNase-seq のピークの高さも同様に Z-Score で計算した。角膜上皮細胞での Z-Score  $> 2.5$  かつその他の 9 細胞での Z-Score  $< 0$  となるピークを検索した結果、2,622 箇所のエンハンサー領域が角膜上皮特異的なエンハンサー領域として抽出された (図 16)。

角膜上皮特異的のマーカである KRT12 や KRT3 の近傍の領域にも角膜上皮特異的なエンハンサー領域が存在していた。KRT12 については TSS から約 7

kb 上流に特異的なピークが見られ (chr17:39,030,125-39,030,755) (図 17)、  
KRT3 については TSS から約 70 kb 下流に特異的なピークが見られた  
(chr12:53,120,260-53,120,912) (図 18)。

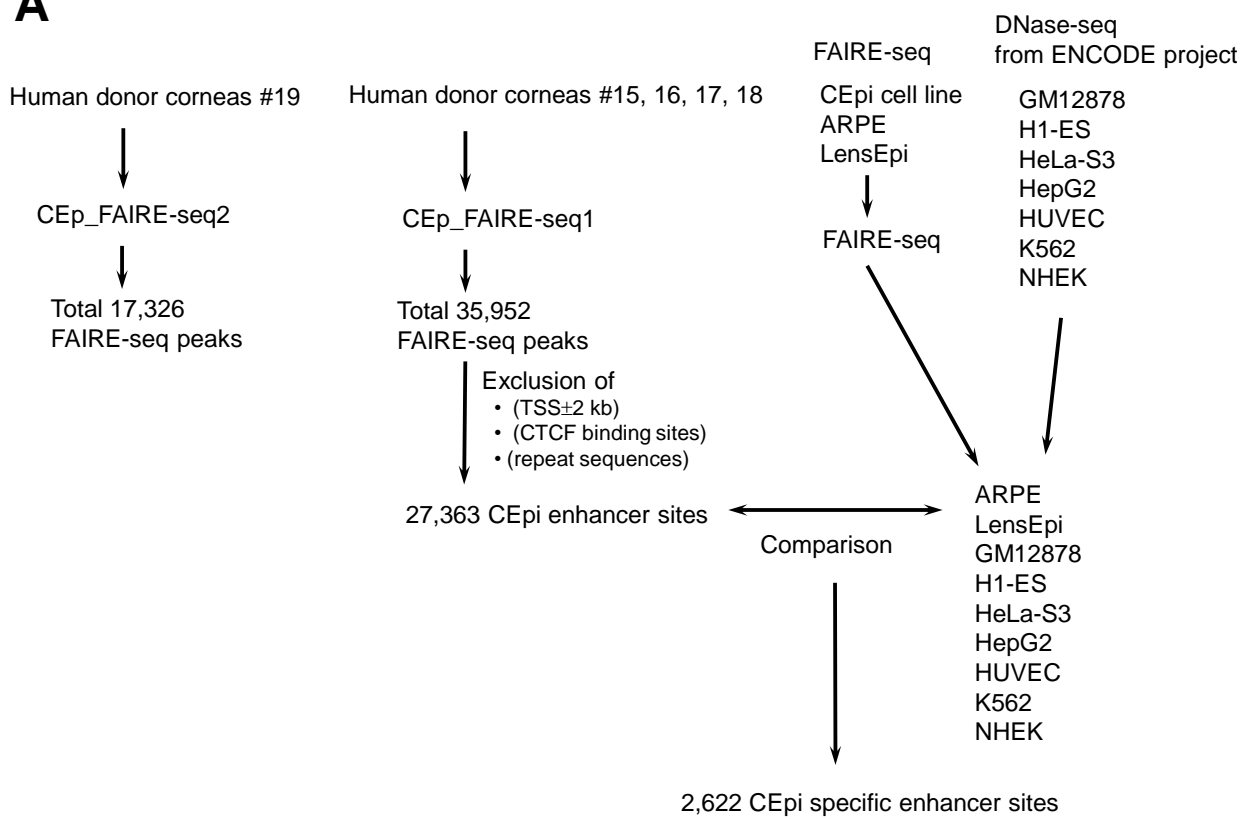
さらに、2,622 箇所の角膜上皮特異的エンハンサー領域の近傍にある 1,589 遺  
伝子 (各エンハンサー領域の遺伝子アノテーションは、最も転写開始点 (TSS:  
Transcription Start Site) の近い遺伝子に割り当てた) について、遺伝子発現  
(FPKM) を比較した。角膜上皮細胞では、他の 9 細胞に比べ、これらの遺伝  
子の発現が高かった ( $p < 2.2e-16$ ) (図 19)。従って今回抽出したエンハンサ  
ー領域は、角膜上皮細胞における遺伝子発現を正に制御していると示唆され  
た。

表7 : FAIRE-seq により同定されたピーク数。

| Sample Name    | Description                        | Read数<br>(PF_cluster) | FAIRE<br>sites |
|----------------|------------------------------------|-----------------------|----------------|
| CEpi_1         | Human donor corneal epithelium     | 39,458,599            | 35,952         |
| CEpi_2         | Human donor corneal epithelium     | 26,437,576            | 17,326         |
| CEpi_cell line | Human corneal epithelial cell line | 28,819,795            | 57,831         |
| ARPE           | Human retinal epithelial cell line | 18,335,403            | 61,730         |
| LensEpi        | Human lens epithelial cell line    | 39,817,214            | 57,403         |

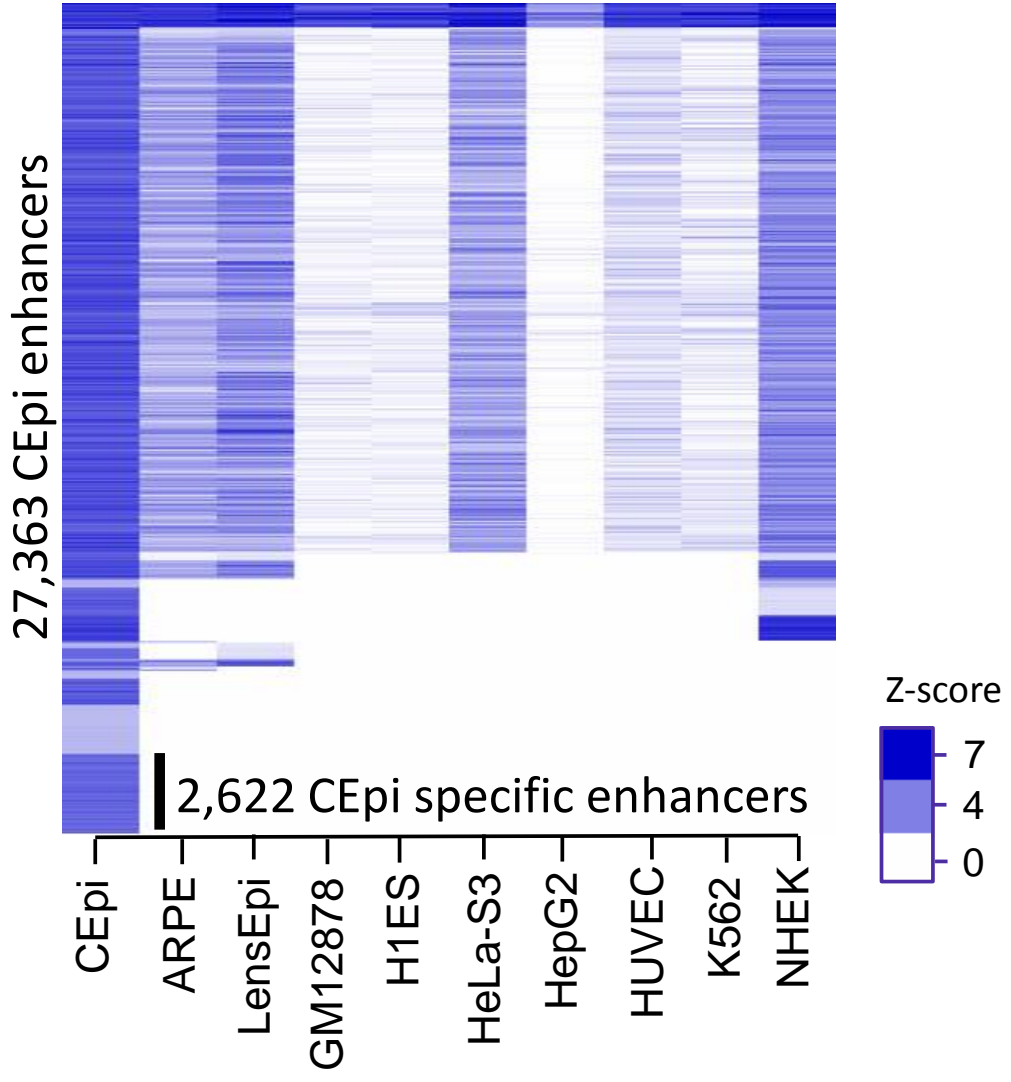
Number of sites detected by MACS software ( $p < 1e-5$ )

**A**





**B**



**図 16:** ヒト角膜上皮特異的エンハンサー領域の同定。(A)ヒト角膜上皮特異的エンハンサー領域の同定の概要。(B) 角膜上皮 (CEpi) エンハンサー領域 27,363 箇所、細胞種別の分布と 2,622 箇所の角膜上皮特異的エンハンサー領域。FAIRE-seq/DNase-seq のピークの高さをヒートマップ (Z-score) で示した。

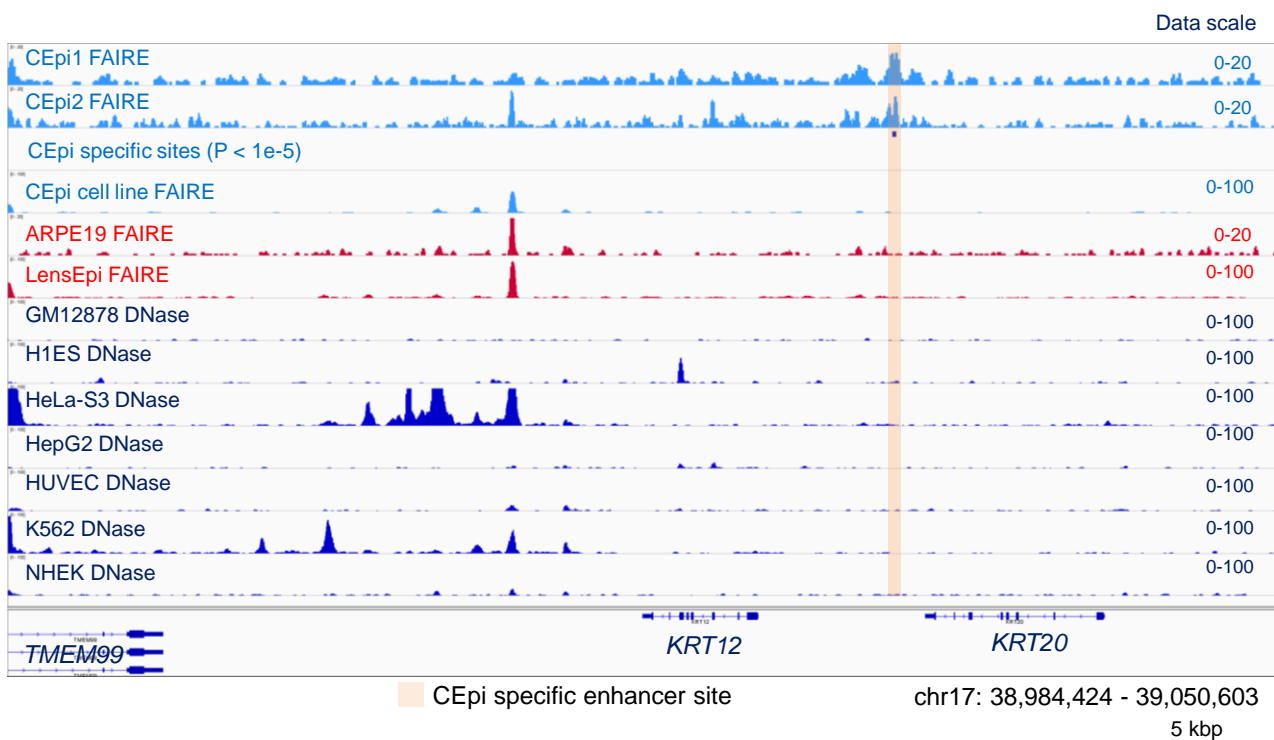


図 17 : KRT12 近傍に存在する角膜上皮特異的エンハンサー領域。

17 番染色体長腕(17q21.2)の KRT12 近傍領域(chr17: 38,984,424 - 39,050,60)を IGV ゲノムブラウザにて示す。角膜上皮特異的エンハンサー領域を橙色網掛で示した。角膜特異的遺伝子 KRT12 の TSS から約 7 kb 上流に角膜上皮特異的エンハンサー領域(橙色網掛)が同定された。

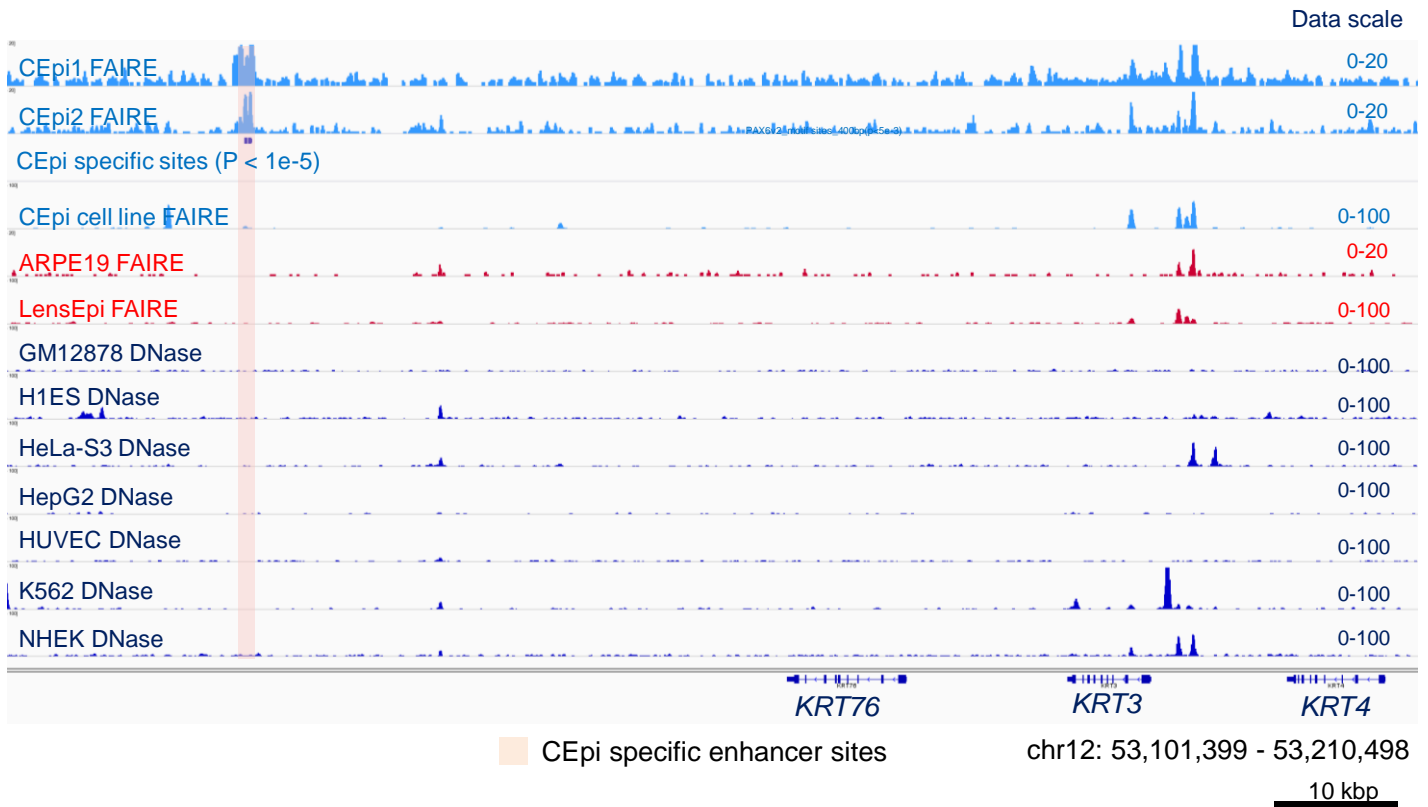


図 18 : KRT3 近傍に存在する角膜上皮特異的エンハンサー領域。

12 番染色体長腕 (12q13.3) の KRT3 近傍領域 (chr12: 53,101,399 - 53,210,498) を IGV ゲノムブラウザにて示す。角膜上皮特異的エンハンサー領域を橙色網掛で示した。角膜特異的遺伝子 KRT3 の TSS から約 70 kb 下流に角膜上皮特異的エンハンサー領域 (橙色網掛) が同定された。

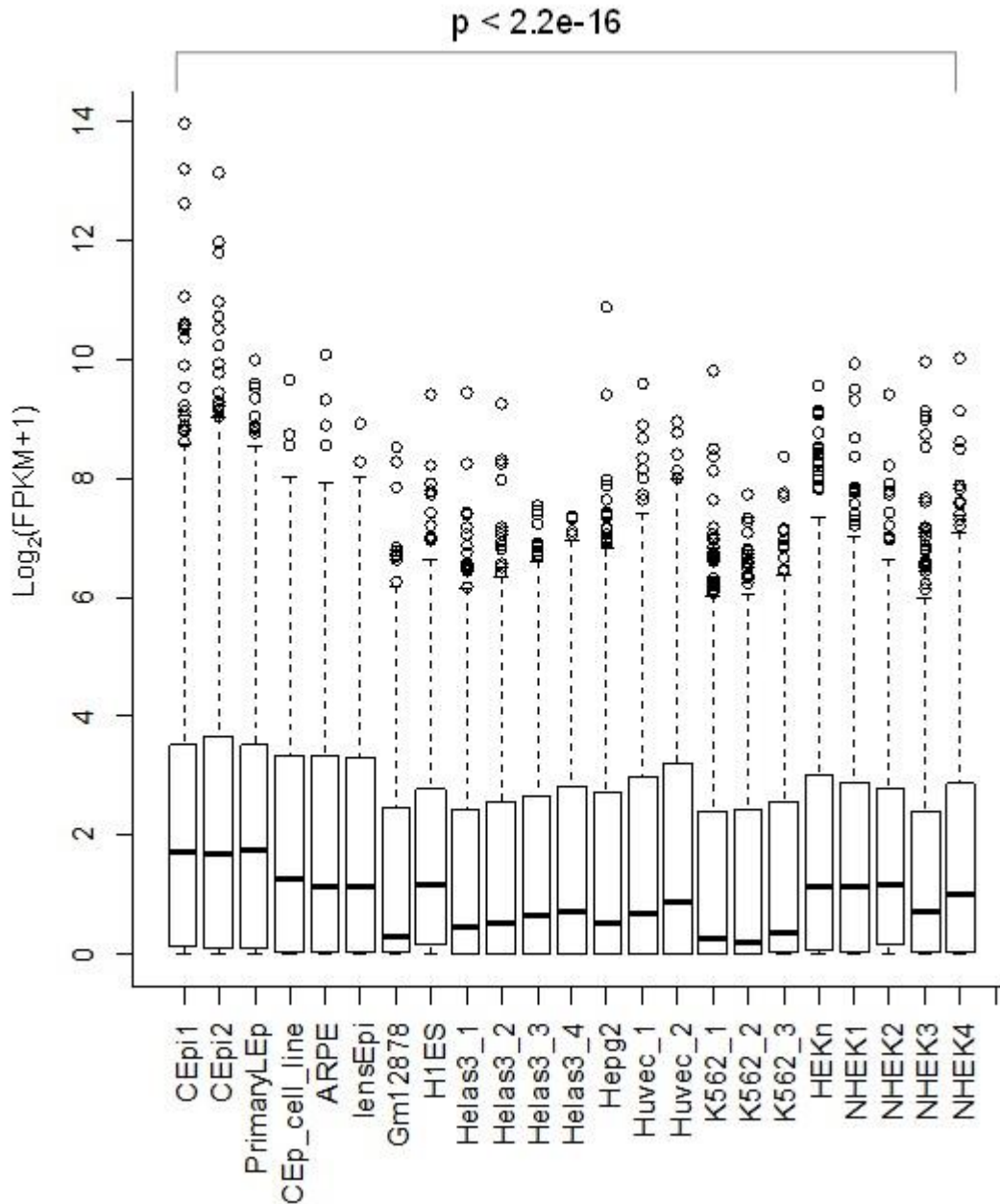


図 19 : 角膜上皮特異的エンハンサー領域の近傍遺伝子の発現量

各細胞における 2,622 箇所の角膜上皮特異的エンハンサー領域近傍の 1,589 遺伝子（各ピーク領域から最も転写開始点（TSS: Transcription Start Site）の近い遺伝子（最大 1M bp 以内））の発現（FPKM）。CEpi1、CEpi2: ドナー角膜上皮細胞、Primary LEp: 初代角膜輪部上皮細胞、CEpi\_cell\_line: 角膜上皮細胞

株。

#### 4C2. 角膜上皮特異的エンハンサー領域のモチーフ解析

2.622 箇所 of 角膜上皮特異的エンハンサー領域について、どのような転写因子のモチーフが濃縮しているのか、既知のモチーフ配列の濃縮を Genomatix と TRAP を用いて解析し、*de novo* モチーフ解析を MODIC のアルゴリズムを用いて行った。

Genomatix では、38 種類の転写因子ファミリーが、Z-Score が 5 以上の濃縮を認めた(表 8A)。発現上昇(mRNA-seq で角膜上皮の全てのサンプルで発現値(FPKM)が皮膚の 2 倍以上かつ角膜上皮の全てのサンプルで FPKM が 1 以上、Exon array で角膜上皮の全てのサンプルで発現値(Raw signal)が皮膚の 2 倍以上かつ角膜上皮の全てのサンプルで Raw signal が 100 以上を共に満たす)を認めていた転写因子(図 15B)のうち、Z-Score が 5 以上の濃縮を認めた転写因子は 8 種類あり、それらは FOSL1 (モチーフ名 : V\$AP1F)、MEIS1 (V\$HASF、V\$PBXC、V\$TALE)、KLF10 (V\$SP1F)、BHLHE41 (V\$HESF)、SMAD3 (V\$SMAD)、PAX6 (V\$PAX6)、KLF3 (V\$E4FF)、KLF7 (V\$E4FF)であった。

TRAP では、38 種類の転写因子モチーフが Corrected P < 1e-20 の濃縮を認めた。これらのモチーフに関連付けられる転写因子のうち、角膜上皮で発現上昇を認めたもの(図 15B)は 7 種類あり、それらは FOSL1 (モチーフ名 : AP1)、PAX6 (PAX6)、NR1D1 (RORA\_2)、MEIS1 (PBX1、HOXA5)、SOX15 (SRY、SOX10、Sox17)、HLF (HLF)、FOXP2 (FOXO3)であった。

MODIC は、*de novo* のモチーフを探索するアルゴリズムであり、本研究ではウィンド

ウサイズを 8 bp、12 bp、16 bp に設定して解析した。検出したモチーフについて、STAMP<sup>74</sup>を用いて JASPAR、TRANSFAC データベースのモチーフで最も近いものを探索した。のべ 60 個の *de novo* モチーフが濃縮しており、重複しているモチーフを除くと、14 個あった。このうち、①2,622 箇所のエンハンサー領域の 5% 以上に濃縮していること、②角膜上皮細胞/ Random genome が 1.5 倍以上、の 2 つの基準を満たすモチーフは、6 種類認めた(表 9C)。これらのうち、角膜上皮で発現上昇を認めていたもの(図 15B)は 4 種類あり、それらは FOSL1(モチーフ名:AP1)、KLF3 (KLF4)、KLF7 (KLF4)、PAX6 (PAX6)であった。

以上より、Genomatix、TRAP、MODIC のいずれかの解析で角膜上皮特異的エンハンサー領域にモチーフの濃縮が認められ、かつ角膜上皮細胞で発現の上昇が認められた転写因子として 13 種類が同定された。これらは PAX6、FOSL1、NR1D1、SOX15、FOXP2、HLF、MEIS1、KLF3、KLF7、KLF10、TGIF1、BHLHE41、SMAD3 であった(表 8D)。

表 8: 角膜上皮特異的エンハンサー領域に濃縮していたモチーフ

(A) Genomatix を用いて抽出したモチーフ

Z-Score (genome) > 5 を満たす上位 38 転写因子ファミリー。Nr. Of Input Seq with Match: 2,622 箇所のエンハンサー領域のうち、少なくとも 1 つ以上のモチーフが含まれる領域の数、Nr. Of Matches in Input: 2,622 箇所のエンハンサー領域に含まれるモチーフの総数、Expected (genome): 同サイズのバックグラウンドゲノムに含まれることが予測されるモチーフの数、Over representation (genome): バックグラウンドゲノムに対する対象領域でのモチーフの出現比すなわち Nr. Of Matches in Input/ Expected (genome)、Z-Score (genome): バックグラウンドゲノムに対する対象領域でのモチーフの出現比を Z-Score で表したものの。

| TF Families | Nr. of Input Seq. with Match | Nr. of Matches in Input | Expected (genome) | Std.dev. | Over representation (genome) | Z-Score (genome) |
|-------------|------------------------------|-------------------------|-------------------|----------|------------------------------|------------------|
| V\$AP1F     | 986                          | 1879                    | 543.68            | 23.3     | 3.46                         | 57.28            |
| V\$AP1R     | 1383                         | 2736                    | 1270.33           | 35.59    | 2.15                         | 41.17            |
| V\$P53F     | 572                          | 986                     | 567.11            | 23.8     | 1.74                         | 17.58            |
| V\$HASF     | 179                          | 187                     | 72.64             | 8.52     | 2.57                         | 13.36            |
| V\$SP1F     | 567                          | 726                     | 482.42            | 21.95    | 1.5                          | 11.07            |
| V\$H1FF     | 267                          | 381                     | 217.36            | 14.74    | 1.75                         | 11.07            |
| V\$PBXC     | 709                          | 865                     | 595.34            | 24.38    | 1.45                         | 11.04            |
| V\$HESF     | 422                          | 617                     | 407.99            | 20.19    | 1.51                         | 10.33            |
| V\$ZF11     | 234                          | 254                     | 135.9             | 11.66    | 1.87                         | 10.09            |
| V\$MYOD     | 647                          | 937                     | 675.2             | 25.96    | 1.39                         | 10.06            |
| V\$ZICF     | 245                          | 273                     | 154.21            | 12.42    | 1.77                         | 9.53             |
| V\$SMAD     | 487                          | 563                     | 383.57            | 19.58    | 1.47                         | 9.14             |
| V\$GRHL     | 427                          | 703                     | 499.49            | 22.34    | 1.41                         | 9.09             |
| V\$AP2F     | 341                          | 539                     | 373.48            | 19.32    | 1.44                         | 8.54             |
| V\$EBOX     | 550                          | 807                     | 599.17            | 24.46    | 1.35                         | 8.48             |
| V\$NEUR     | 578                          | 727                     | 531.6             | 23.04    | 1.37                         | 8.46             |
| V\$PAX6     | 823                          | 1079                    | 836.71            | 28.9     | 1.29                         | 8.37             |
| V\$CP2F     | 405                          | 461                     | 312.78            | 17.68    | 1.47                         | 8.36             |
| V\$NFKB     | 445                          | 600                     | 432.65            | 20.79    | 1.39                         | 8.03             |
| V\$TALE     | 629                          | 754                     | 565.58            | 23.77    | 1.33                         | 7.91             |
| V\$BRAC     | 660                          | 832                     | 651.52            | 25.51    | 1.28                         | 7.06             |
| V\$MIZ1     | 161                          | 170                     | 100.09            | 10       | 1.7                          | 6.94             |
| V\$RBP2     | 137                          | 144                     | 81.38             | 9.02     | 1.77                         | 6.89             |
| V\$PAX5     | 659                          | 841                     | 663.21            | 25.73    | 1.27                         | 6.89             |
| V\$PAX3     | 446                          | 492                     | 365.46            | 19.11    | 1.35                         | 6.6              |
| V\$LTFM     | 217                          | 227                     | 150.61            | 12.27    | 1.51                         | 6.18             |
| V\$AP4R     | 186                          | 228                     | 152.34            | 12.34    | 1.5                          | 6.09             |
| V\$FXRE     | 189                          | 223                     | 150.39            | 12.26    | 1.48                         | 5.88             |
| V\$HICF     | 286                          | 312                     | 224.5             | 14.98    | 1.39                         | 5.81             |
| V\$MEF3     | 271                          | 281                     | 199.02            | 14.1     | 1.41                         | 5.78             |
| O\$XCPE     | 126                          | 132                     | 80.36             | 8.96     | 1.64                         | 5.7              |
| V\$ZFHX     | 607                          | 723                     | 590.3             | 24.28    | 1.22                         | 5.44             |
| V\$WHNF     | 101                          | 106                     | 62.94             | 7.93     | 1.68                         | 5.36             |
| V\$ZFXF     | 132                          | 145                     | 94.32             | 9.71     | 1.54                         | 5.17             |
| V\$YBXF     | 203                          | 218                     | 153.7             | 12.4     | 1.42                         | 5.15             |
| V\$KLFS     | 931                          | 1388                    | 1209.91           | 34.74    | 1.15                         | 5.11             |
| V\$E4FF     | 220                          | 247                     | 178.37            | 13.35    | 1.38                         | 5.1              |
| V\$PAX9     | 74                           | 82                      | 46.75             | 6.84     | 1.75                         | 5.08             |







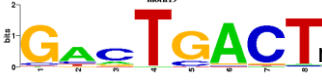

**(B) TRAP** を用いて抽出したモチーフ

corrected p-value < 1e-20 を満たす上位 38 転写因子を示す。転写因子配列は JAPSPAR を用いた。

| #/ Rank | Corrected_P | Matrix_ID | Matrix_name   |
|---------|-------------|-----------|---------------|
| 1       | 0           | MA0099.2  | <b>AP1</b>    |
| 2       | 7.83E-245   | MA0075.1  | <b>Prrx2</b>  |
| 3       | 4.43E-133   | MA0150.1  | <b>NFE2L2</b> |
| 4       | 1.66E-91    | MA0067.1  | <b>Pax2</b>   |
| 5       | 8.32E-81    | MA0125.1  | <b>Nobox</b>  |
| 6       | 4.52E-77    | MA0069.1  | <b>Pax6</b>   |
| 7       | 3.35E-76    | MA0072.1  | <b>RORA_2</b> |
| 8       | 5.48E-76    | MA0132.1  | <b>Pdx1</b>   |
| 9       | 3.88E-60    | MA0151.1  | <b>ARID3A</b> |
| 10      | 4.41E-60    | MA0153.1  | <b>HNF1B</b>  |
| 11      | 6.81E-59    | MA0124.1  | <b>NKX3-1</b> |
| 12      | 3.52E-56    | MA0135.1  | <b>Lhx3</b>   |
| 13      | 1.22E-55    | MA0025.1  | <b>NFIL3</b>  |
| 14      | 3.35E-53    | MA0070.1  | <b>PBX1</b>   |
| 15      | 5.70E-53    | MA0084.1  | <b>SRY</b>    |
| 16      | 8.17E-51    | MA0038.1  | <b>Gfi</b>    |
| 17      | 6.11E-50    | MA0009.1  | <b>T</b>      |
| 18      | 6.16E-50    | MA0148.1  | <b>FOXA1</b>  |
| 19      | 1.04E-48    | MA0158.1  | <b>HOXA5</b>  |
| 20      | 3.70E-40    | MA0087.1  | <b>Sox5</b>   |
| 21      | 3.42E-39    | MA0063.1  | <b>Nkx2-5</b> |
| 22      | 7.64E-39    | MA0142.1  | <b>Pou5f1</b> |
| 23      | 3.62E-38    | MA0047.2  | <b>Foxa2</b>  |
| 24      | 1.59E-37    | MA0442.1  | <b>SOX10</b>  |
| 25      | 2.54E-37    | MA0078.1  | <b>Sox17</b>  |
| 26      | 1.81E-36    | MA0043.1  | <b>HLF</b>    |
| 27      | 7.20E-33    | MA0031.1  | <b>FOXD1</b>  |
| 28      | 2.09E-32    | MA0027.1  | <b>En1</b>    |
| 29      | 1.69E-31    | MA0040.1  | <b>Foxq1</b>  |
| 30      | 9.94E-30    | MA0102.2  | <b>CEBPA</b>  |
| 31      | 7.19E-27    | MA0033.1  | <b>FOXL1</b>  |
| 32      | 1.35E-26    | MA0157.1  | <b>FOXO3</b>  |
| 33      | 1.78E-26    | MA0103.1  | <b>ZEB1</b>   |
| 34      | 1.70E-25    | MA0090.1  | <b>TEAD1</b>  |
| 35      | 5.00E-23    | MA0032.1  | <b>FOXC1</b>  |
| 36      | 1.06E-22    | MA0030.1  | <b>FOXF2</b>  |
| 37      | 2.42E-21    | MA0029.1  | <b>Evi1</b>   |
| 38      | 1.24E-20    | MA0143.1  | <b>Sox2</b>   |

### (C) MODIC を用いて抽出したモチーフ

*de novo* motifs: 今回見つかったモチーフ。Database: JASPAR、TRANSFAC にて STAMP にて照合した転写因子名。CEpi: 角膜上皮特異的エンハンサー領域での各モチーフの出現割合。Random (Promoter): バックグラウンド(プロモーター)における各モチーフの出現割合。Random (genome): バックグラウンド(ゲノム)における各モチーフの出現割合。CEpi/ Random (genome): 各転写因子のバックグラウンド(ゲノム)に対する角膜上皮特異的エンハンサー領域における出現比。

| de novo motifs  | Name                     | Database      | CEn    | CEpi   | Random (promoter) | Random (genome) | CEpi/Random (genome) |
|---|--------------------------|---------------|--------|--------|-------------------|-----------------|----------------------|
|  | CEp_12bp_<br>motif1.pwm  | <b>AP1</b>    | 0.0428 | 0.1110 | 0.0327            | 0.0353          | <b>3.1422</b>        |
|  | CEp_12bp_<br>motif11.pwm | <b>AP1</b>    | 0.0614 | 0.1773 | 0.0549            | 0.0631          | <b>2.8105</b>        |
|  | CEp_12bp_<br>motif14.pwm | <b>NFE2L2</b> | 0.0178 | 0.0645 | 0.0229            | 0.0234          | <b>2.7568</b>        |
|  | CEp_16bp_<br>motif1.pwm  | <b>KLF4</b>   | 0.0450 | 0.0751 | 0.1177            | 0.0466          | <b>1.6109</b>        |
|  | CEp_8bp_<br>motif19.pwm  | <b>AP1</b>    | 0.0885 | 0.1819 | 0.1089            | 0.1060          | <b>1.7162</b>        |
|  | CEp_8bp_<br>motif4.pwm   | <b>PAX6</b>   | 0.0314 | 0.0641 | 0.0233            | 0.0357          | <b>1.7958</b>        |

(D) 59 個の角膜上皮高発現転写因子における、2,622 箇所の角膜上皮特異的エンハンサー領域に濃縮していたモチーフ

Genomatix、TRAP、MODIC にて濃縮している転写因子 (Z-Score>5 (Genomatix)、corrected p-value <1e-20 (TRAP)、CEpi/ Background (genome) >1.5 (MODIC)) を赤色で示した。

| gene symbol | Genomatix | TRAP | MODIC |
|-------------|-----------|------|-------|
| PAX6        | Red       |      |       |
| FOSL1       | Red       |      |       |
| ELF3        |           |      |       |
| ZBED2       |           |      |       |
| BHLHE41     | Red       |      |       |
| EAF2        |           |      |       |
| ENO1        |           |      |       |
| IRF1        |           |      |       |
| SMAD3       | Red       |      |       |
| KLF3, 7     | Red       |      | Red   |
| TGIF1       |           |      |       |
| FOXP2       |           | Red  |       |
| MAPK10      |           |      |       |
| NR4A2       |           | Red  |       |
| SOX15       |           | Red  |       |
| NKX3-1      |           |      |       |
| KLF10       | Red       |      |       |
| HLF         |           | Red  |       |
| TSC22D2     |           |      |       |
| PIR         |           |      |       |
| RB1         |           |      |       |
| PIM1        |           |      |       |
| ID1         |           |      |       |
| TRIM16      |           |      |       |
| ELAVL2      |           |      |       |
| NLK         |           |      |       |
| MEIS1       | Red       |      |       |
| IRF6        |           |      |       |
| MBNL3       |           |      |       |
| NR1D1       |           | Red  |       |
| PPIL1       |           |      |       |
| ZNF326      |           |      |       |
| THOC3       |           |      |       |
| LRRFIP1     |           |      |       |
| TRETF1      |           |      |       |
| PARP4       |           |      |       |
| TNNI2       |           |      |       |
| ZNF597      |           |      |       |
| SERTAD3     |           |      |       |
| PDLIM1      |           |      |       |
| L3MBTL3     |           |      |       |
| HMG3        |           |      |       |
| ZNF687      |           |      |       |
| PLAGL2      |           |      |       |
| APLP2       |           |      |       |
| ZNF440      |           |      |       |
| SRA1        |           |      |       |
| BNC1        |           |      |       |
| RBM4        |           |      |       |
| ZNF827      |           |      |       |
| MS1         |           |      |       |
| TCEB1       |           |      |       |
| PPRC1       |           |      |       |
| ZNF593      |           |      |       |
| MORF4L1     |           |      |       |
| ZNF207      |           |      |       |
| TAF5L       |           |      |       |
| DDX19B      |           |      |       |

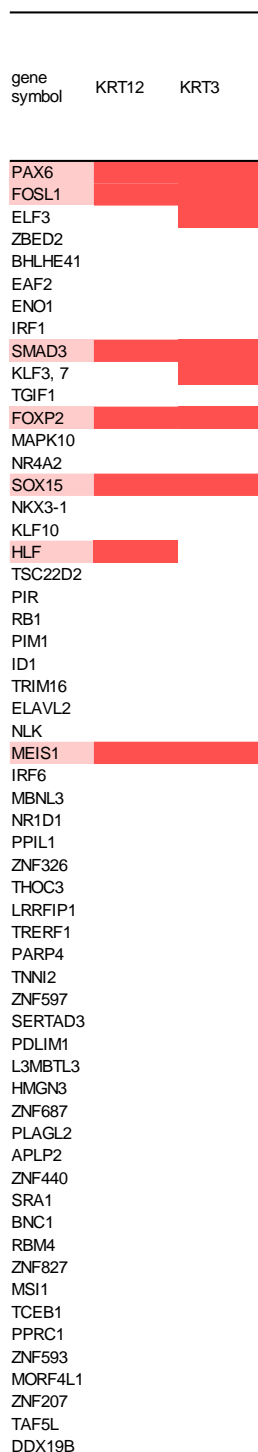
### 4C3. KRT12近傍およびKRT3近傍の角膜上皮特異的エンハンサー領域のモチーフ

#### 解析

角膜上皮特異マーカー遺伝子 KRT12 近傍の角膜上皮特異的エンハンサー領域 (chr17:39,030,125-39,030,755) (図 17) にどのようなモチーフが存在するか検索した。その結果、Genomatix によって、PAX6、FOSL1、SOX15、FOXP2、HLF、MEIS1、SMAD3 を、TRAP によって HLF、PAX6、FOSL1、SOX15 を、MODIC によって FOSL1 と PAX6 のモチーフをそれぞれ認めた。すなわち、KRT12 近傍の角膜上皮特異的エンハンサー領域にモチーフを認め、角膜上皮で発現が上昇する転写因子は、PAX6、FOSL1、SOX15、FOXP2、HLF、MEIS1、SMAD3 であった(表 9)。また、KRT3 近傍の角膜上皮特異的エンハンサー領域には、PAX6、FOSL1、ELF3、SMAD3、KLF3、KLF7、FOXP2、SOX15、MEIS1 のモチーフが存在した(表 9)。

表 9. 59 個の角膜上皮高発現転写因子における、KRT12 近傍の角膜上皮特異的エンハンサーに存在するモチーフ。

KRT12 の TSS より約 7 kb 上流の角膜上皮特異的エンハンサー領域 (chr17:39030451-39030644) に存在するモチーフと、KRT3 の TSS より約 70 kb 下流の角膜上皮特異的エンハンサー領域 (chr12:53120610-53120895) に存在するモチーフ赤で示す。



## V. 考察

### 5A. 角膜上皮特異的遺伝子の同定

Exon array を用いた、ヒトドナー角膜上皮とヒト正常組織・培養細胞 81 サンプルの網羅的な遺伝子発現解析により、角膜上皮特異的なマーカー遺伝子を探索した。

角膜上皮特異的遺伝子として、①角膜上皮の全てのサンプルで発現値 (Raw signal) が皮膚の 10 倍以上、かつ②角膜上皮の全てのサンプルで発現値 (Raw signal) がヒト正常組織・培養細胞 81 サンプルの中央値の 10 倍以上を満たす遺伝子は、52 遺伝子見つかった。中でも、角膜での特異性が特に高い遺伝子として、ヒト正常組織・培養細胞 81 サンプルのうち角膜上皮の発現値 (Raw signal) の 0.1 倍以上の組織・細胞の数が 0 または 1 である遺伝子は、KRT12、PAX6、KRT3、KRT24 であった (表 5、図 9)。

KRT12 の発現は角膜以外の他組織 (今回調べた 81 個のヒト正常組織・培養細胞) では認めず (図 9A)、最も特異性の高いマーカー遺伝子であることが示された。

一方、KRT3 の発現は舌に認められた (図 9B)。KRT3 は口腔粘膜に発現しているとの報告がある<sup>80</sup>。しかし、他のヒト正常組織 80 種類では KRT3 の発現を認めず、KRT12 と組み合わせることで有用なマーカーであると考えられた。

PAX6 は脳で発現を認めた (図 9C) が、KRT3、KRT12 と組み合わせることで、有用なマーカーとなり得ると考えられた。KRT24 は、食道、子宮頸部、胎盤でわずかに発現を認めた (図 9D)。

KRT24 はこれまで、舌、胎盤、大腸での発現が報告されている<sup>81,82</sup>。その他 52 個の中に含まれている遺伝子としては、ALDH3A1<sup>83</sup> は角膜で高発現であること<sup>78</sup>、ELF3 は KRT12 のプロモーターを活性化すること<sup>84</sup>、TGFB1 は角膜ジストロフィの原因遺伝子として<sup>85-87</sup>、WNT7A は角膜上皮細胞の特異性の維持に関わる可能性が報告されている<sup>47</sup>。

## 5B. 初代培養角膜輪部上皮細胞と角膜上皮細胞株の遺伝子発現解析

無血清・無フィーダー培養による、初代培養角膜輪部上皮細胞について、網羅的な発現解析を行った。その結果、初代培養細胞は *in vivo* の角膜上皮細胞に近い遺伝子発現プロファイルを持つことが確認できた。初代培養細胞において発現 (FPKM>1) が確認できたのは、52 個の角膜上皮特異的遺伝子のうち 47 個であり (図 13)、KRT3、KRT12、PAX6 のいずれも発現が確認された (図 12)。

一方、広く用いられている SV40 large T 抗原不死化角膜上皮細胞株は<sup>51,54</sup>、角膜特異的 64k Da サイトケラチン (KRT3) を発現しているとされているが、今回、通常の培養条件 DMEM/ F12 FBS 10% でも、初代培養と同じ無血清・無フィーダーの培養条件<sup>31,32</sup> でも、KRT3、KRT12、PAX6 の発現を確認できなかった (図 12、14)。細胞株において発現 (FPKM>1) が確認できたのは、52 個の角膜上皮特異的遺伝子のうち 30 個に留まった (図 13)。Greco らもこの細胞株における KRT3 の発現は大幅に低下しており KRT12、PAX6 はほとんど発現していないことを報告している<sup>88</sup>。

したがって、初代培養角膜輪部上皮細胞を KRT3、KRT12 をはじめとする角膜上皮細胞の特異性を規定する転写制御の解析に使うことは可能と考えられた。特に、*in vitro* でのノックダウン実験には KRT12 を発現している培養上皮細胞が不可欠である。また、細胞数を要するゲノムワイドなエピゲノム解析が容易になると考えられた。

一方、不死化ヒト角膜上皮細胞株では KRT3、KRT12、PAX6 の発現が低下しており、角膜上皮の特異性を規定する転写因子を明らかにする実験には適さないと考えられた。

また、52 個の角膜上皮特異的遺伝子のうち、細胞株と初代培養細胞株で差があった遺伝子は、KRT12、PAX6、KRT3、KRT24、ALDH3A1、MAB21L1、TMPRSS11D、CRTAC1、TMPRSS4、HRASLS2、ADH7、HORMAD1、GDPD2、RASSF6、APOBEC3A、WNT7A、CPAMD8、AGR2 の 18 遺伝子であった。ALDH3A1 は角膜上皮細胞での発現が報告されている<sup>89</sup>。また、WNT7A は角膜上皮細胞の特異性の維持に関わるとの報告がある<sup>47</sup>。52 遺伝子の中でこれら 18 遺伝子は、角膜上皮細胞の特異性の維持に関わっている特に重要な角膜上皮特異的遺伝子である可能性がある。

初代培養細胞の問題点として、KRT3 は FPKM が 2.17 と低発現であり、KRT3 は発現はしているものの、KRT12 や PAX6 に比較して決して十分とは言えない。上皮細胞の培養において、マウス 3T3feeder 細胞を用いることは広く行われているが、われわれは臨床応用を考慮して動物由来の組織の排除のため、feeder free で 5~6 層の重層化培養上皮シートを得られている<sup>31,32</sup> (図



11) 。 KRT3 の発現に関しては、feeder 細胞を用いた培養法における遺伝子発現を検討の余地があると考えられる。

### 5C. ヒトドナー角膜上皮細胞の FAIRE-seq

本研究では、ヒトドナー角膜上皮細胞の FAIRE-seq を行った。これまで、FAIRE-seq や DNase-seq は、細胞株または初代培養細胞を用いたもので、*in vivo* での検討はマウスでの報告であった。今回、細胞株 (CEpi cell line、ARPE19、LenEpi) では、57,000~62,000 箇所のヌクレオソーム・フリー領域が同定されたのに対して、*in vivo* ヒトドナー角膜上皮細胞では、それよりやや少なく 17,000~36,000 箇所のヌクレオソーム・フリー領域が同定された (表 7) 。これは、図 17、図 18 から分かるように、*in vivo* ヒトドナー角膜上皮細胞では、細胞株に比べてシグナル/ ノイズ比が高いことによるものと思われた。

角膜上皮の *in vivo* FAIRE が比較的良好に行えた理由は、角膜上皮細胞の採取が容易で、他の細胞や繊維組織等の混入が少ないこと、ホルムアルデヒドが速やかに浸透しクロマチンを均一に固定できること、細胞数が多く取れた (20 万~100 万細胞/ 1 ドナー強角膜片) ことが考えられる。

### 5D. KRT3、KRT12 近傍のエンハンサー領域の同定

今まで、KRT12 のプロモーター領域は調べられてきた<sup>90</sup>が、KRT3、KRT12 近傍<sup>40</sup>のエンハンサー領域は明らかでなかった。今回、FAIRE-seq によりゲノムワイドにヌクレオソーム・フリー領域を調べることにより、KRT3、KRT12 近傍に存在するエンハンサー

領域を同定した。

特に、角膜特異的遺伝子 **KRT12** クラスター領域に存在するエンハンサー領域 (chr17:39,030,125-39,030,755) に関しては、**KRT12** がこのエンハンサー領域から最も近傍 (7 kb) にあり (図 17)、また、近傍に **KRT12** 以外に発現している遺伝子は存在しないことから、このエンハンサー領域が **KRT12** の発現を正に制御するエンハンサー領域であることが示唆された。

また、角膜特異的遺伝子 **KRT3** クラスター領域に存在するエンハンサー領域 (chr12:53,120,260-53,120,912) に関しては、この角膜上皮特異的エンハンサー領域から近傍には **KRT77**、**KRT76**、**KRT1** があるが、これらの遺伝子は角膜上皮に発現しておらず、このエンハンサー領域から約 70 kb に角膜上皮特異的遺伝子 **KRT3** がある (図 18)。このエンハンサー領域から上流・下流 約 150 kb に **CTCF** 結合領域を認め、この **CTCF** 結合領域で囲まれた約 300 kb 以内に、**KRT3** 以外に角膜上皮で発現している遺伝子は、認めなかった。

今回同定したこれらのエンハンサー領域が **KRT3**、**KRT12** の発現を本当に制御しているかはさらなる研究が必要である。より直接的には、3C (Chromosome conformation capture) 法、4C (Circularized Chromosome Conformation Capture) 法にて、エンハンサー領域と **KRT3**、**KRT12** の TSS 領域の空間的近接関係を解析することや、**CRISPR/Cas9** システムを用い、エンハンサー領域を欠損させることで **KRT3**、**KRT12** の発現が抑制されるかを調べる必要がある。

## 5E. エンハンサー領域と遺伝子発現の相関

図 19 に示すように、角膜上皮特異的な 2,622 箇所のエンハンサー領域近傍の遺伝子発現は、角膜上皮において高かった。これは、既報でも同様である<sup>57</sup>。

ただ、エンハンサー領域を、最も近傍に TSS がある遺伝子に、アノテーションするというやり方（方法の 2H10. の項参照）が、必ずしも正確ではない。たとえば、図 18 に示すように、KRT3 領域に存在する角膜上皮特異的エンハンサー領域は、最も近傍の TSS を持つ遺伝子は KRT77 であるが、この遺伝子は角膜上皮で発現していない。それぞれのエンハンサー領域を直接調節を受ける遺伝子にアノテーションできるような、より優れた方法の開発が待たれる。Hi-C 法など、クロマチンの空間的近接関係の網羅的な解析が有用だと考えられるが、現状のプロトコルでは、 $10^7 \sim 10^8$  個程度の細胞数を必要である。したがって、臨床組織など細胞数の限られた材料を扱う実験に適用するためには、さらなるプロトコルの改良が必要である。

## 5F. 角膜上皮細胞の特異性維持に関わる 7 転写因子

本研究では、ヒト角膜上皮細胞の網羅的発現解析およびクロマチン構造解析を行った。網羅的発現解析の結果、59 転写因子が皮膚と比較して角膜上皮細胞で高発現であった（図 15B）。ゲノムワイドな角膜上皮特異的エンハンサー領域のモチーフ解析と、KRT12 近傍のエンハンサー領域のモチーフから、59 転写因子の中で 7 転写因子（PAX6、FOSL1、SOX15、FOXP2、HLF、MEIS1、SMAD3）が角膜上皮の特異性維持に関わる転写因子候補として見出された（表 8D、表 9）。

7 転写因子の中で角膜上皮細胞の発生に関わる転写因子としては、PAX6 が知ら

れている<sup>91</sup>が、それ以外は不明である。また、PAX6 に関しては、生前の眼形成と、生後・成人における角膜上皮細胞の細胞系譜の維持(KRT12 の発現の維持)における重要な役割が示唆されている。PAX6 は、眼球発生のマスターレギュレーターとして発見されている。PAX6 遺伝子は、Drosophila からヒトやマウスを含む哺乳類まで構造的、機能的に保存されている<sup>92,93</sup>。PAX6 の変異は、ヒトにおいては無虹彩症、マウスにおいては small eye (Sey) の表現型を示す<sup>92</sup>。

マウスにおいて、KRT12 の発現は、生後 4 日目より角膜上皮が重層化を始めるときに弱く発現を始め、眼瞼が開く前に発現が増強し、角膜上皮が重層化を完成させた後、成期において高発現を保つ<sup>94,95</sup>。KRT12 の発現は、角膜上皮の分化に依存しており、PAX6 は成体マウスの角膜上皮で高発現であることから、成体マウスの KRT12 の発現を増強している可能性がある<sup>96</sup>。マウスにおいて、他組織では成期になると PAX6 の発現は失われるが、眼表面の分化した重層上皮(角膜上皮・角膜輪部上皮・結膜上皮)では依然として高発現であり、角膜や結膜の遺伝子制御に直接かかわっていることが示唆されている<sup>96</sup>。ウサギの角膜上皮の細胞株において、ルシフェラーゼアッセイにより KRT12 のプロモーター活性を PAX6 が転写制御していることが報告され、PAX6 が角膜上皮の細胞系譜の維持に重要な役割を持つことが示唆された<sup>90</sup>。

また、Stevens-Johnson 症候群や無虹彩症などの角膜輪部機能不全の患者において、眼表面の PAX6 の発現が減少していることが報告されている<sup>48,97</sup>。また、このような眼表面疾患の進行過程において共通して認める眼表面の扁平上皮化成において、KRT12 から KRT10 への発現の移行が観察されている<sup>97,98</sup>。現在まで、眼表面における PAX6 の減少と扁平上皮化成の関連について、因果関係を証明するまでには至っ

ていない。さらに、Ouyang H らは、表皮ケラチノサイトに PAX6 を導入すると、KRT3、KRT12 の発現が上昇する、と報告した<sup>47</sup>。

一方、他の 6 転写因子 FOSL1、SOX15、FOXP2、HLF、MEIS1、SMAD3 に関しては、角膜の特異性維持、例えば KRT12 の発現に関わっているとの報告は無い。ノックアウトマウスや、ヒト遺伝性疾患においても、他 6 転写因子が関わる角膜疾患は報告されていない。しかし、これら 6 転写因子は角膜上皮において皮膚と比較して高発現であること、角膜上皮特異的エンハンサー領域 2,622 箇所にもモチーフの濃縮が認められること、角膜上皮特異的 KRT12 近傍のエンハンサー領域にもモチーフが存在することから、角膜上皮細胞の特異性を規定している可能性があると考えられた。中でも HLF は、KRT3 近傍のエンハンサーに存在せず、KRT12 のエンハンサーに存在していること、結膜と比較して角膜上皮において発現が高いこと(未発表データ)から、KRT12 の特異性を規定している可能性が高い。今回の研究から推測される、転写制御ネットワークを図 19 に示した。

## 5G. 今後の展開

本研究において、角膜上皮細胞の特異性の維持に関わる可能性のある転写因子を同定した。今回同定した転写因子は、角膜上皮細胞において発現が上昇していること、角膜上皮特異的エンハンサー領域の転写因子結合モチーフから絞り込んだものである。実際にこれらの転写因子が特異性維持に関わっているかどうか、さらなる研究が必要である。

具体的には、今後の展開として、①KRT12 をはじめとして角膜上皮細胞の特異的マ

一カー遺伝子の発現の確認された初代培養角膜輪部上皮細胞を用いて、sh (pLKO.1)を用いた転写因子のノックダウンにより、KRT12 が抑制されるかどうか解析する、②初代培養表皮角化細胞に7転写因子をレンチウイルスベクターを用いて導入し、KRT12 が発現するかどうか解析する、ことを計画している。

また、本研究においては角膜上皮細胞を皮膚と比較しているが、角膜上皮と同様の非角化重層扁平上皮で KRT3 陽性の口腔粘膜・PAX6 陽性の結膜との比較をすることで、PAX6 以外の角膜上皮細胞の特異性を規定する転写因子により迫っていけるものと考えている。

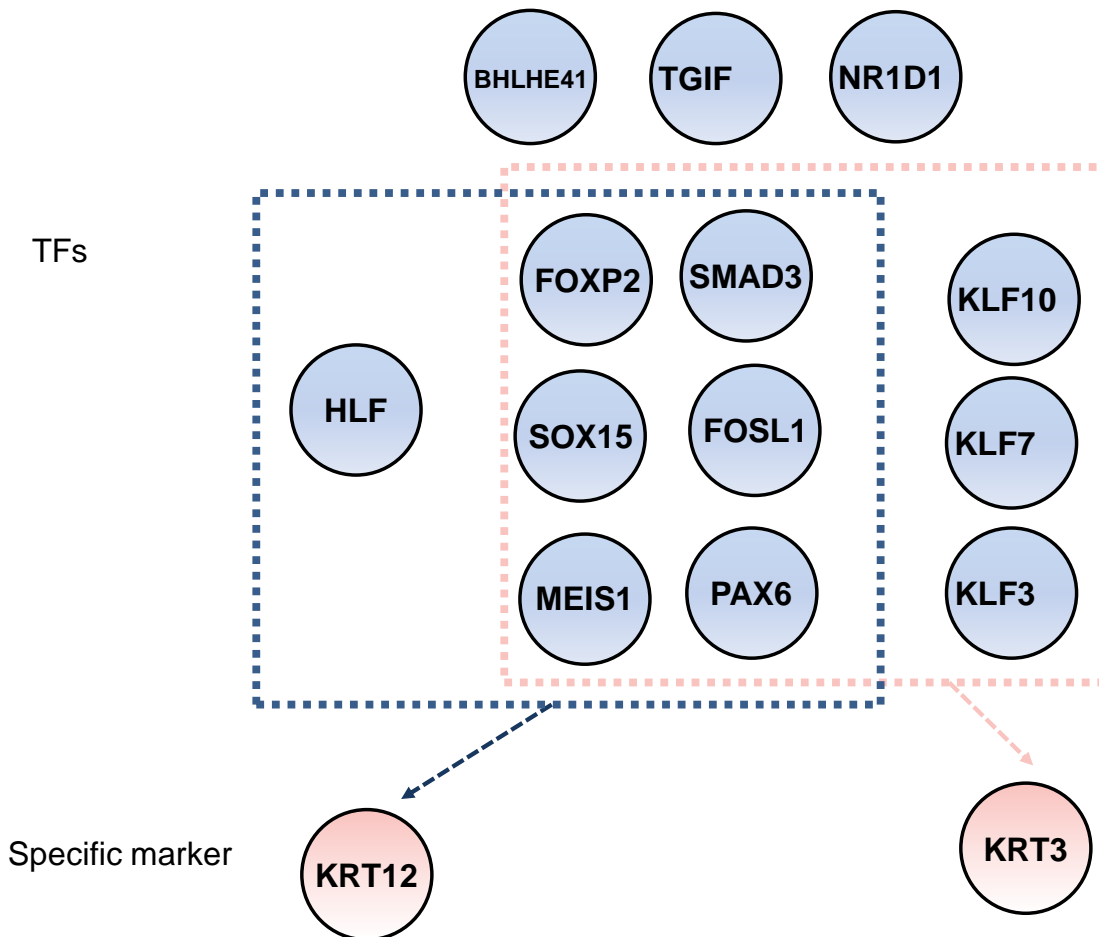


図 19. 網羅的発現解析とモチーフ解析の統合解析から推測した角膜上皮特異性を規定する転写因子

角膜上皮特異的エンハンサー2,622箇所濃縮しておりかつ角膜上皮細胞で発現が特徴的に上昇していた13転写因子を示す。青四角囲みは、KRT12近傍のエンハンサー領域にモチーフが存在する7転写因子を示し、赤四角囲みは、KRT3近傍のエンハンサー領域にモチーフが存在する9転写因子を示す。

## VI. まとめ

本研究では、ヒト角膜上皮細胞の Exon array と mRNA-seq による網羅的発現解析と FAIRE-seq によるクロマチン構造解析を行うことにより、角膜上皮細胞の特異的遺伝子と特異性を規定する転写因子の同定を目的とした。

まず、Exon array にて角膜上皮特異的マーカー遺伝子の探索を行い、52 個の角膜上皮特異的マーカー遺伝子候補を見出した。既報のマーカー遺伝子 KRT12 や KRT3 の角膜上皮細胞（組織）における特異的な発現が示された。中でも KRT12 は、今回調べた 81 のヒト正常組織・培養細胞に発現しているサンプルは認められず、最も特異性の高いマーカー遺伝子であることが示された。

次に、無血清・無フィーダー培養による初代培養角膜輪部上皮細胞と、角膜上皮細胞株におけるマーカー遺伝子の発現を mRNA-seq を用いて探索した。初代培養細胞では 52 個のマーカー遺伝子の多くが発現している一方、角膜上皮細胞株ではその発現が限定されることが示された。

さらに、FAIRE-seq により、ゲノムワイドに 35,952 箇所のヌクレオソーム・フリー領域を同定した。ヒト 9 細胞と比較することにより、2,622 個の角膜上皮特異的エンハンサー領域を同定した。KRT3、KRT12 近傍にも角膜上皮特異的エンハンサー領域が存在した。2,622 箇所の近傍の遺伝子は、角膜上皮において発現が高く、角膜上皮特異的エンハンサー領域と遺伝子の発現の間に正の相関が存在した。2,622 箇所のエンハンサー領域に濃縮している特徴的な転写因子の結合配列（モチーフ）を解析したところ、モチーフが濃縮しておりかつ



角膜上皮で発現が上昇している転写因子が 13 個同定された。このうち、PAX6、FOSL1、SOX15、FOXP2、HLF、SMAD3、MEIS1 の 7 転写因子は KRT12 近傍のエンハンサー領域にモチーフが確認されたことから、角膜上皮の特異性を規定する可能性が高い転写因子と考えられた。

## 「謝辞」

本研究の指導委託教官である東京大学先端科学技術研究センターゲノムサイエンス分野、油谷浩幸教授には、本研究の遂行にあたり、多大なるご高配とご指導、ご鞭撻を賜りました。

本研究の遂行にあたり、東京大学大学院医学系研究科外科学専攻眼科学教室、新家眞前々教授、天野史郎前教授、山上聡准教授、加藤聡准教授、相原一前講師、臼井智彦講師、横尾誠一助教、大道幸子研究員には、多大なるご高配とご指導、ご鞭撻を賜りました。

本研究の解析にあたり、東京大学先端科学技術研究センターゲノムサイエンス分野、仲木竜研究員、山本省吾博士研究員、上田宏生博士研究員、藤田隆教研究員、後藤健吾研究員に多大なるご指導、ご協力を賜りました。

本研究の遂行にあたり、東京大学先端科学技術研究センターゲノムサイエンス分野、永江玄太助教、関元昭博士研究員、野村征太郎博士研究員に多大なるご指導、ご鞭撻を賜りました。

本研究全般において、東京大学先端科学技術研究センターゲノムサイエンス分野、辰野健二博士研究員、野中綾博士研究員、岡部篤史博士研究員、神尾明日香博士研究員、林 玲匡研究員、串田夏樹研究員、東京大学大学院医学系研究科外科学専攻眼科学教室、島伸之助教、木枕光木子研究員、愛新覚羅愛研究員には、数々のご教示、ご協力を頂きました。

本研究の試料作製にあたり、東京大学先端科学技術研究センターゲノムサイエンス分野、椎名香織技術員、目黒裕子技術研究員、川辺さおり技術研究員、藤中恭子技術研究員に多大なるご協力を賜りました。

宮田眼科病院 宮田和典先生、森洋斉先生からは、研究用ヒトドナー強角膜片を一部ご提供頂きました。厚く御礼申し上げます。

本研究で、ご指導、ご鞭撻、ご高配、ご協力賜りました上記の方々、および東京大学先端科学技術研究センターゲノムサイエンス分野、東京大学大学院医学系研究科眼科学教室の皆様、に、深謝申し上げます次第です。

## 「引用文献」

1. Schermer, A., Galvin, S. & Sun, T. T. Differentiation-related expression of a major 64K corneal keratin in vivo and in culture suggests limbal location of corneal epithelial stem cells. *J. Cell Biol.* **103**, 49–62 (1986).
2. Cotsarelis, G., Cheng, S. Z., Dong, G., Sun, T. T. & Lavker, R. M. Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: implications on epithelial stem cells. *Cell* **57**, 201–9 (1989).
3. Oie, Y. & Nishida, K. Regenerative medicine for the cornea. *Biomed Res Int* **2013**, 428247 (2013).
4. Pellegrini, G., Rama, P., Mavilio, F. & De Luca, M. Epithelial stem cells in corneal regeneration and epidermal gene therapy. *J. Pathol.* **217**, 217–28 (2009).
5. Watanabe, K., Nishida, K., Yamato, M., Umemoto, T., Sumide, T., Yamamoto, K., Maeda, N., Watanabe, H., Okano, T. & Tano, Y. Human limbal epithelium contains side population cells expressing the ATP-binding cassette transporter ABCG2. *FEBS Lett.* **565**, 6–10 (2004).
6. Schlötzer-Schrehardt, U. & Kruse, F. Identification and characterization of limbal stem cells. *Experimental Eye Research* **81**, 247264 (2005).
7. Hayashi, R., Yamato, M., Saito, T., Oshima, T., Okano, T., Tano, Y. & Nishida, K. Enrichment of corneal epithelial stem/progenitor cells using cell surface markers, integrin alpha6 and CD71. *Biochem. Biophys. Res. Commun.* **367**, 256–63 (2008).
8. Hayashi, R., Yamato, M., Sugiyama, H., Sumide, T., Yang, J., Okano, T., Tano, Y. & Nishida, K. N-Cadherin is expressed by putative stem/progenitor cells and melanocytes in the human limbal epithelial stem cell niche. *Stem Cells* **25**, 289–96 (2007).
9. Whitcher, J. P., Srinivasan, M. & Upadhyay, M. P. Corneal blindness: a global perspective. *Bull. World Health Organ.* **79**, 214–21 (2001).

10. McAlinden, C. Corneal refractive surgery: past to present. *Clin Exp Optom* **95**, 386–98 (2012).
11. Colin, J. Ganciclovir ophthalmic gel, 0.15%: a valuable tool for treating ocular herpes. *Clin Ophthalmol* **1**, 441–53 (2007).
12. Tan, D. T., Dart, J. K., Holland, E. J. & Kinoshita, S. Corneal transplantation. *Lancet* **379**, 1749–61 (2012).
13. Ahmad, S, Figueiredo, F & Lako, M. Corneal epithelial stem cells: characterization, culture and transplantation. (2006). at <http://www.futuremedicine.com/doi/abs/10.2217/17460751.1.1.29>
14. Sareen, D., Saghizadeh, M., Ornelas, L., Winkler, M. A., Narwani, K., Sahabian, A., Funari, V. A., Tang, J., Spurka, L., Punj, V., Maguen, E., Rabinowitz, Y. S., Svendsen, C. N. & Ljubimov, A. V. Differentiation of human limbal-derived induced pluripotent stem cells into limbal-like epithelium. *Stem Cells Transl Med* **3**, 1002–12 (2014).
15. Chee, K. Y., Kicic, A. & Wiffen, S. J. Limbal stem cells: the search for a marker. *Clin. Experiment. Ophthalmol.* **34**, 64–73 (2006).
16. Tsubota, K., Satake, Y., Kaido, M., Shinozaki, N., Shimmura, S., Bissen-Miyajima, H. & Shimazaki, J. Treatment of Severe Ocular-Surface Disorders with Corneal Epithelial Stem-Cell Transplantation. *New England Journal of Medicine* (1999). doi:10.1056/NEJM199906033402201
17. Kenyon, K. R. & Tseng, S. C. Limbal autograft transplantation for ocular surface disorders. *Ophthalmology* **96**, 709–22; discussion 722–3 (1989).
18. Pellegrini, G., Traverso, C. E., Franzi, A. T., Zingirian, M., Cancedda, R. & De Luca, M. Long-term restoration of damaged corneal surfaces with autologous cultivated corneal epithelium. *Lancet* **349**, 990–3 (1997).
19. Nakamura, T., Endo, K.-I., Cooper, L. J., Fullwood, N. J., Tanifuji, N., Tsuzuki, M., Koizumi, N., Inatomi, T., Sano, Y. & Kinoshita, S. The successful culture

- and autologous transplantation of rabbit oral mucosal epithelial cells on amniotic membrane. *Invest. Ophthalmol. Vis. Sci.* **44**, 106–16 (2003).
20. Nishida, K., Yamato, M., Hayashida, Y., Watanabe, K., Yamamoto, K., Adachi, E., Nagai, S., Kikuchi, A., Maeda, N., Watanabe, H., Okano, T. & Tano, Y. Corneal reconstruction with tissue-engineered cell sheets composed of autologous oral mucosal epithelium. *N. Engl. J. Med.* **351**, 1187–96 (2004).
21. Kolli, S., Ahmad, S., Mudhar, H. S., Meeny, A., Lako, M. & Figueiredo, F. C. Successful application of ex vivo expanded human autologous oral mucosal epithelium for the treatment of total bilateral limbal stem cell deficiency. *Stem Cells* (2014). doi:10.1002/stem.1694
22. Satake, Y., Higa, K., Tsubota, K. & Shimazaki, J. Long-term outcome of cultivated oral mucosal epithelial sheet transplantation in treatment of total limbal stem cell deficiency. *Ophthalmology* **118**, 1524–30 (2011).
23. Satake, Y., Yamaguchi, T., Hirayama, M., Higa, K., Shimazaki-Den, S., Dogru, M., Kawakita, T., Kawashima, M., Shimmura, S., Tsubota, K. & Shimazaki, J. Ocular surface reconstruction by cultivated epithelial sheet transplantation. *Cornea* **33 Suppl 11**, S42–6 (2014).
24. Tanioka, H., Kawasaki, S., Yamasaki, K., Ang, L., Koizumi, N., Nakamura, T., Yokoi, N., Komuro, A., Inatomi, T. & Kinoshita, S. Establishment of a Cultivated Human Conjunctival Epithelium as an Alternative Tissue Source for Autologous Corneal Epithelial Transplantation. *Investigative Ophthalmology & Visual Science* (2006). doi:10.1167/iovs.06-0293
25. Usui, T., Nakagawa, S., Yokoo, S., Mimura, T., Yamagami, S. & Amano, S. Bilateral limbal stem cell deficiency with chromosomal translocation of 3p and 9p. *Jpn. J. Ophthalmol.* **54**, 357–8 (2010).
26. Hayashi, R., Ishikawa, Y., Ito, M., Kageyama, T., Takashiba, K., Fujioka, T., Tsujikawa, M., Miyoshi, H., Yamato, M., Nakamura, Y. & Nishida, K. Generation of

corneal epithelial cells from induced pluripotent stem cells derived from human dermal fibroblast and corneal limbal epithelium. *PLoS ONE* **7**, e45435 (2012).

27. O'Callaghan, A. & Daniels, J. Concise review: limbal epithelial stem cell therapy: controversies and challenges. *Stem cells (Dayton, Ohio)* **29**, 1923–32 (2011).

28. Hewitt, K. J., Shamis, Y., Carlson, M. W., Aberdam, E., Aberdam, D. & Garlick, J. A. Three-dimensional epithelial tissues generated from human embryonic stem cells. *Tissue Eng Part A* **15**, 3417–26 (2009).

29. Cooper DKC. The Promise of Transplanting Animal Organs into Humans. *New York: Oxford University Press* 7 17 (2000).

30. U.S. Department of Health and Human Services Food and Drug Administration (FDA) Center for Biologics Evaluation and Research (CBER). Information and Recommendations for Physicians Involved in the Co-culture of Human Embryos with Nonhuman Animal Cells. *U.S. Food and Drug Administration: Rockville MD* (2002).

31. Yokoo, S., Yamagami, S., Usui, T., Amano, S. & Araie, M. Human corneal epithelial equivalents for ocular surface reconstruction in a complete serum-free culture system without unknown factors. *Invest. Ophthalmol. Vis. Sci.* **49**, 2438–43 (2008).

32. Nakagawa, S., Usui, T., Yokoo, S., Omichi, S., Kimakura, M., Mori, Y., Miyata, K., Aihara, M., Amano, S. & Araie, M. Toxicity evaluation of antiglaucoma drugs using stratified human cultivated corneal epithelial sheets. *Invest. Ophthalmol. Vis. Sci.* **53**, 5154–60 (2012).

33. Odom, D. T., Dowell, R. D., Jacobsen, E. S., Nekludova, L., Rolfe, P. A., Danford, T. W., Gifford, D. K., Fraenkel, E., Bell, G. I. & Young, R. A. Core transcriptional regulatory circuitry in human hepatocytes. *Mol. Syst. Biol.* **2**, 2006.0017 (2006).

34. Odom, D. T., Zizlsperger, N., Gordon, D. B., Bell, G. W., Rinaldi, N. J., Murray, H. L., Volkert, T. L., Schreiber, J., Rolfe, P. A., Gifford, D. K., Fraenkel, E., Bell, G. I. & Young, R. A. Control of pancreas and liver gene expression by HNF

transcription factors. *Science* **303**, 1378–81 (2004).

35. Sekiya, S. & Suzuki, A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475**, 390–393 (2011).

36. Chaloin-Dufau, C., Sun, T. T. & Dhouailly, D. Appearance of the keratin pair K3/K12 during embryonic and adult corneal epithelial differentiation in the chick and in the rabbit. *Cell Differ. Dev.* **32**, 97–108 (1990).

37. Moll, R, Franke, WW, Schiller, DL, Geiger, B & Krepler, R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *cell* (1982). at

<<http://www.sciencedirect.com/science/article/pii/0092867482904007>>

38. Moll, R., Franke, W. W., Schiller, D. L., Geiger, B. & Krepler, R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell* **31**, 11–24 (1982).

39. Schweizer, J., Bowden, P. E., Coulombe, P. A., Langbein, L., Lane, E. B., Magin, T. M., Maltais, L., Omary, M. B., Parry, D. A., Rogers, M. A. & Wright, M. W. New consensus nomenclature for mammalian keratins. *J. Cell Biol.* **174**, 169–74 (2006).

40. Rogers, MA. Keratins and keratin diseases. *eLS* (2003).

doi:10.1038/npg.els.0005905

41. Converse, R, Candace, W. & Sun, TT. Cornea-specific expression of K12 keratin during mouse development. *Current eye ...* (1993). at

<<http://informahealthcare.com/doi/abs/10.3109/02713689309029222>>

42. Wei, Z. G., Wu, R. L., Lavker, R. M. & Sun, T. T. In vitro growth and differentiation of rabbit bulbar, fornix, and palpebral conjunctival epithelia. Implications on conjunctival epithelial transdifferentiation and stem cells. *Invest. Ophthalmol. Vis. Sci.* **34**, 1814–28 (1993).

43. Chaloin-Dufau, C., Pavitt, I., Delorme, P. & Dhouailly, D. Identification of keratins 3 and 12 in corneal epithelium of vertebrates. *Epithelial Cell Biol* **2**, 120–5



(1993).

44. Hesse, M., Zimek, A., Weber, K. & Magin, T. Comprehensive analysis of keratin gene clusters in humans and rodents. *European Journal of Cell Biology* **83**, 1926 (2004).

45. Irvine, AD, Corden, LD, Swensson, O & Swensson, B. Mutations in cornea-specific keratin K3 or K12 genes cause Meesmann's corneal dystrophy. *Nature* ... (1997). at <<http://www.nature.com/ng/journal/v16/n2/abs/ng0697-184.html>>

46. Kao, WW, Liu, CY, Converse, RL & Shiraishi, A. Keratin 12-deficient mice have fragile corneal epithelia. ... *ophthalmology & visual* ... (1996). at <<http://www.iovs.org/content/37/13/2572.short>>

47. Ouyang, H., Xue, Y., Lin, Y., Zhang, X., Xi, L., Patel, S., Cai, H., Luo, J., Zhang, M., Zhang, M., Yang, Y., Li, G., Li, H., Jiang, W., Yeh, E., *et al.* WNT7A and PAX6 define corneal epithelium homeostasis and pathogenesis. *Nature* **511**, 358–61 (2014).

48. Li, W., Chen, Y.-T. T., Hayashida, Y., Blanco, G., Kheirkah, A., He, H., Chen, S.-Y. Y., Liu, C.-Y. Y. & Tseng, S. C. Down-regulation of Pax6 is associated with abnormal differentiation of corneal epithelial cells in severe ocular surface diseases. *J. Pathol.* **214**, 114–22 (2008).

49. Wang, C., Gong, B., Bushel, P., Thierry-Mieg, J., Thierry-Mieg, D., Xu, J., Fang, H., Hong, H., Shen, J., Su, Z., Meehan, J., Li, X., Yang, L., Li, H., Łabaj, P., *et al.* The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nature Biotechnology* (2014). doi:10.1038/nbt.3001

50. Kratz, A. & Carninci, P. The devil in the details of RNA-seq. *Nature Biotechnology* **32**, 882–884 (2014).

51. Araki-Sasaki, K., Ohashi, Y., Sasabe, T., Hayashi, K., Watanabe, H., Tano, Y. & Handa, H. An SV40-immortalized human corneal epithelial cell line and its characterization. *Invest. Ophthalmol. Vis. Sci.* **36**, 614–21 (1995).

52. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annual review of biochemistry* **81**, (2012).
53. Qu, H. & Fang, X. A brief review on the Human Encyclopedia of DNA Elements (ENCODE) project. *Genomics Proteomics Bioinformatics* **11**, 135–41 (2013).
54. Kahn, C. R., Young, E., Lee, I. H. & Rhim, J. S. Human corneal epithelial primary cultures and cell lines with extended life span: in vitro model for ocular studies. *Invest. Ophthalmol. Vis. Sci.* **34**, 3429–41 (1993).
55. Heinäniemi, M., Nykter, M., Kramer, R., Wienecke-Baldacchino, A., Sinkkonen, L., Zhou, J., Kreisberg, R., Kauffman, S., Huang, S. & Shmulevich, I. Gene-pair expression signatures reveal lineage control. *Nature methods* **10**, 577–83 (2013).
56. Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., Kuehn, M. S., Taylor, C. M., Neph, S., Koch, C. M., Asthana, S., *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
57. Song, L., Zhang, Z., Grasfeder, L. L., Boyle, A. P., Giresi, P. G., Lee, B.-K. K., Sheffield, N. C., Gräf, S., Huss, M., Keefe, D., Liu, Z., London, D., McDaniel, R. M., Shibata, Y., Showers, K. A., *et al.* Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.* **21**, 1757–67 (2011).
58. Stergachis, A. B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S. L., Vernot, B., Cheng, J. B., Thurman, R. E., Sandstrom, R., Haugen, E., Heimfeld, S., Murry, C. E., Akey, J. M. & Stamatoyannopoulos, J. A. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell* **154**, 888–903 (2013).
59. Ong, C.-T. & Corces, V. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nature reviews. Genetics* **12**, 283–93 (2011).
60. Bulger, M. & Groudine, M. Functional and mechanistic diversity of distal

transcription enhancers. *Cell* **144**, 327–39 (2011).

61. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Molecular cell* **49**, 825–37 (2013).
62. Heintzman, N., Hon, G., Hawkins, R., Kheradpour, P., Stark, A., Harp, L., Ye, Z., Lee, L., Stuart, R., Ching, C., Ching, K., Antosiewicz-Bourget, J., Liu, H., Zhang, X., Green, R., *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–12 (2009).
63. Giresi, P., Kim, J., McDaniell, R., Iyer, V. & Lieb, J. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome research* **17**, 877–85 (2007).
64. Simon, J. M., Giresi, P. G., Davis, I. J. & Lieb, J. D. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat Protoc* **7**, 256–67 (2012).
65. Voss, T., Schiltz, R., Sung, M.-H., Yen, P., Stamatoyannopoulos, J., Biddie, S., Johnson, T., Miranda, T., John, S. & Hager, G. Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. *Cell* **146**, 544–54 (2011).
66. Waki, H., Nakamura, M., Yamauchi, T., Wakabayashi, K., Yu, J., Hirose-Yotsuya, L., Take, K., Sun, W., Iwabu, M., Okada-Iwabu, M., Fujita, T., Aoyama, T., Tsutsumi, S., Ueki, K., Kodama, T., Sakai, J., Aburatani, H. & Kadowaki, T. Global mapping of cell type-specific open chromatin by FAIRE-seq reveals the regulatory role of the NFI family in adipocyte differentiation. *PLoS Genet.* **7**, e1002311 (2011).
67. Parkinson, N. J., Maslau, S., Ferneyhough, B., Zhang, G., Gregory, L., Buck, D., Ragoussis, J., Ponting, C. P. & Fischer, M. D. Preparation of high-quality next-generation sequencing libraries from picogram quantities of target DNA. *Genome Res.* **22**, 125–33 (2012).
68. Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., Bernstein, B.,

- Nusbaum, C., Myers, R., Brown, M., Li, W. & Liu, X. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137 (2008).
69. Thorvaldsdóttir, H., Robinson, J. & Mesirov, J. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* **14**, 178–92 (2013).
70. Barrera, L., Li, Z., Smith, A., Arden, K., Cavenee, W., Zhang, M., Green, R. & Ren, B. Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome research* **18**, 46–59 (2008).
71. Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. & Werner, T. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics (Oxford, England)* **21**, 2933–42 (2005).
72. Thomas-Chollier, M., Hufton, A., Heinig, M., O’Keeffe, S., Masri, N., Roider, H., Manke, T. & Vingron, M. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nature protocols* **6**, 1860–9 (2011).
73. Nakaki, R., Kang, J. & Tateno, M. A novel ab initio identification system of transcriptional regulation motifs in genome DNA sequences based on direct comparison scheme of signal/noise distributions. *Nucleic Acids Res.* **40**, 8835–48 (2012).
74. Mahony, S. & Benos, P. V. STAMP: a web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res.* **35**, W253–8 (2007).
75. Boyle, A. P., Song, L., Lee, B.-K. K., London, D., Keefe, D., Birney, E., Iyer, V. R., Crawford, G. E. & Furey, T. S. High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Res.* **21**, 456–64 (2011).
76. Sandelin, A., Alkema, W., Engström, P., Wasserman, W. & Lenhard, B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic acids research* **32**, D91–4 (2004).

77. Mathelier, A, Zhao, X, Zhang, AW & Parcy, F. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic acids* ... (2013). doi:10.1093/nar/gkt997
78. Matys, V., Kel-Margoulis, O. V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., Voss, N., Stegmaier, P., Lewicki-Potapov, B., Saxel, H., Kel, A. E. & Wingender, E. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34**, D108–10 (2006).
79. Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A., Kel-Margoulis, O., Kloos, D.-U., Land, S., Lewicki-Potapov, B., Michael, H., Münch, R., *et al.* TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic acids research* **31**, 374–8 (2003).
80. Shimazaki, J., Higa, K., Kato, N. & Satake, Y. Barrier function of cultivated limbal and oral mucosal epithelial cell sheets. *Invest. Ophthalmol. Vis. Sci.* **50**, 5672–80 (2009).
81. Rogers, M. A., Winter, H., Langbein, L., Bleiler, R. & Schweizer, J. The human type I keratin gene family: characterization of new hair follicle specific members and evaluation of the chromosome 17q21.2 gene domain. *Differentiation* **72**, 527–40 (2004).
82. Bragulla, H. & Homberger, D. Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *Journal of anatomy* **214**, 516–59 (2009).
83. PAPPAS, A., ESTEY, T., MANZER, R., BROWN, D. & VASILIOU, V. Human aldehyde dehydrogenase 3A1 (ALDH3A1): biochemical characterization and immunohistochemical localization in the cornea. *Biochemical Journal* **376**, 615 (2003).
84. Yoshida, N., Yoshida, S., Araie, M., Handa, H. & Nabeshima, Y. Ets family transcription factor ESE-1 is expressed in corneal epithelial cells and is involved in their

differentiation. *Mechanisms of Development* **97**, (2000).

85. Stewart, H., Ridgway, A., Dixon, M., Bonshek, R., Parveen, R. & Black, G. Heterogeneity in granular corneal dystrophy: Identification of three causative mutations in the TGFBI (BIGH3) gene—Lessons for corneal amyloidogenesis. *Hum. Mutat.* (1999). doi:10.1002/(SICI)1098-1004(1999)14:2<126::AID-HUMU4>3.0.CO;2-W

86. Stewart, H., Black, G. C., Donnai, D., Bonshek, R. E., McCarthy, J., Morgan, S., Dixon, M. J. & Ridgway, A. A. A mutation within exon 14 of the TGFBI (BIGH3) gene on chromosome 5q31 causes an asymmetric, late-onset form of lattice corneal dystrophy. *Ophthalmology* **106**, 964–970 (1999).

87. Dighiero, P., Niel, F., Ellies, P., D’Hermies, F. & Savoldelli, M. Histologic phenotype–genotype correlation of corneal dystrophies associated with eight distinct mutations in the TGFBI gene. *Ophthalmology* (2001). at <<http://www.sciencedirect.com/science/article/pii/S016164200000662X>>

88. Greco, D., Vellonen, K.-S. S., Turner, H. C., Häkli, M., Tervo, T., Auvinen, P., Wolosin, J. M. & Urtti, A. Gene expression analysis in SV-40 immortalized human corneal epithelial cells cultured with an air-liquid interface. *Mol. Vis.* **16**, 2109–20 (2010).

89. Kays, W. T. & Piatigorsky, J. Aldehyde dehydrogenase class 3 expression: identification of a cornea-preferred gene promoter in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13594–9 (1997).

90. Shiraishi, A., Converse, R. L., Liu, C. Y., Zhou, F., Kao, C. W. & Kao, W. W. Identification of the cornea-specific keratin 12 promoter by in vivo particle-mediated gene transfer. *Invest. Ophthalmol. Vis. Sci.* **39**, 2554–61 (1998).

91. Collinson, J. M., Quinn, J., Hill, R. & West, J. The roles of Pax6 in the cornea, retina, and olfactory epithelium of the developing mouse embryo. *Developmental Biology* **255**, 303–312 (2003).

92. Strachan, T. & Read, A. PAX genes. *Current opinion in genetics &*

*development* **4**, 427–38 (1994).

93. Noll, M. Evolution and role of Pax genes. *Current opinion in genetics & development* **3**, 595–605 (1993).

94. Liu, C., Zhu, G., Westerhausen-Larson, A., Converse, R., Kao, C., Sun, T. & Kao, W. Cornea-specific expression of K12 keratin during mouse development. *Current eye research* **12**, 963–74 (1993).

95. Liu, C. Y., Zhu, G., Converse, R., Kao, C. W., Nakamura, H., Tseng, S. C., Mui, M. M., Seyer, J., Justice, M. J. & Stech, M. E. Characterization and chromosomal localization of the cornea-specific murine keratin gene Krt1.12. *J. Biol. Chem.* **269**, 24627–36 (1994).

96. Koroma, B., Yang, J. & Sundin, O. The Pax-6 homeobox gene is expressed throughout the corneal and conjunctival epithelia. *Investigative ophthalmology & visual science* **38**, 108–20 (1997).

97. Chen, Y. T., Chen, F. Y., Vijmasi, T., Stephens, D. N., Gallup, M. & McNamara, N. A. Pax6 downregulation mediates abnormal lineage commitment of the ocular surface epithelium in aqueous-deficient dry eye disease. *PLoS ONE* **8**, e77286 (2013).

98. Tseng, S., Maumenee, A., Stark, W., Maumenee, I., Jensen, A., Green, W. & Kenyon, K. Topical retinoid treatment for various dry-eye disorders. *Ophthalmology* **92**, 717–27 (1985).

—