Single-cell transcriptomics uncover conserved molecular mechanisms and functional diversification in multilayered epithelia

Candice Merle[1], Mathilde Huyghe[1], Erica Kimber[1], Marisa M. Faraldo[1], Sophie Pantalacci[2], Marie Semon[2], Silvia Fre[1]* & Robin P. Journot[1]*†

† Lead contact

*Corresponding authors Email: silvia.fre@curie.fr (SF) robinjournot5@gmail.com (RJ)

[1] Institut Curie, PSL University, Sorbonne Université, CNRS UMR3215, Inserm U934, Genetics and Developmental Biology, 75005 Paris, France

2 Laboratory of Biology and Modeling of the Cell, Ecole Normale Supérieure de Lyon, CNRS, Inserm, UCBL, 69007, Lyon, France

Keywords

Epithelial stratification, p63-Notch axis, Single-cell transcriptomics, Comparative genomics, Gene regulatory networks

Abstract

Multilayered epithelia, including skin, cervix, thymus, and prostate, arise from all three germ layers and perform diverse physiological functions, yet share a conserved layered architecture in which a supportive basal layer underlies differentiated suprabasal cells. To determine whether this common organization reflects a shared genetic program or tissue-specific regulation, we assembled a comprehensive single-cell atlas of 14 murine Multilayered epithelia.

We identified a conserved p63-Notch axis that universally governs basal cell identity maintenance and suprabasal commitment. The p63-centered transcriptional program is shared, while Notch signaling recruits context-specific transcriptional modules to generate functional diversity. The layered architecture of epithelia spatially and temporally decouples conserved basal from context-specific suprabasal networks, compartmentalizing functional innovation across multilayered epithelia.

Comparative genomics traces the p63-Notch axis to the root of vertebrates over 500 million years ago, showing that it was repeatedly co-opted through the incorporation of newly evolved genes in the suprabasal domain to generate novel epithelial functions. Our work establishes that evolution of multilayered epithelia operates through modular, compartmentalized diversification of a conserved scaffold of a conserved architectural scaffold, providing a unifying principle for how complex tissues achieve functional innovation while preserving structural integrity.

Introduction

Multilayered epithelia are among the most adaptable tissues in vertebrates, forming protective barriers, secretory glands, and immunological interfaces across organs such as the skin, mammary gland, and thymus[1–3]. Arising from all three embryonic germ layers, ectoderm, endoderm, and mesoderm[4–11], these tissues perform diverse physiological roles, yet they share a remarkably similar architecture: a basal compartment of progenitors supports suprabasal layers that differentiate into specialized lineages. This paradox of conserved architecture despite functional divergence raises a fundamental question: do multilayered epithelia share a universal molecular program, or does each organ deploy its own regulatory mechanisms to build a similar structure?

Work in the epidermis has described complementary regulatory roles for p63, a member of the p53 family that maintains basal progenitors[12–16], and Notch signaling, which promotes suprabasal commitment[17–22]**.** Similar patterns of p63 expression in basal cells have been reported in other multilayered epithelia, including the trachea, bladder, mammary gland, and thymus[13,23–25], although functional evidence outside the epidermis remains limited. Recent observations from our lab in glandular epithelia suggest that the p63–Notch axis may operate more broadly[26,27], but whether this axis represents a universal regulatory principle of epithelial stratification remains unknown.

Multilayered epithelia first arose in vertebrates with epidermal stratification[28], coinciding with the emergence of p63[29,30]. This raises the possibility that a conserved molecular framework underlies stratification, yet direct evidence across tissues is lacking.

Here we address this gap by constructing a single-cell transcriptomic atlas of 14 mouse multilayered epithelia spanning diverse germ-layer origins and functions. By integrating these data with lineage tracing, genetic perturbations, and gene regulatory network (GRN) inference, we reveal the existence of a universal spatial and temporal decoupling of two regulatory networks: a conserved basal program that maintains progenitor identity, and a modular suprabasal program that drives tissue-specific functional diversity.

Comparative genomics demonstrates that this regulatory compartmentalization is reflected in evolution, as suprabasal cells preferentially expressing newly evolved, lineage-specific genes. This indicates that the architectural layering partitions the deployment of evolutionary innovations. Our work thus defines a unifying framework for understanding multilayered epithelia organogenesis, revealing how architectural robustness can be balanced with functional diversity.

Results

A single-cell transcriptomic atlas of 14 mouse multilayered epithelia

To systematically map transcriptional diversity across multilayered epithelia, we compiled a comprehensive single-cell atlas of over 500,000 cells from 14 mouse organs, covering stratified and pseudo-stratified epithelia of diverse germ-layer origins, ages, sexes, and genetic backgrounds (Figure S1A-D). This atlas integrates 146 publicly available single-cell RNA sequencing (scRNA-seq) datasets from the epidermis[31–33], cornea[34–36], oral mucosa[37–39], thymus[40–42], mammary gland (MG)[43–46], vagina[47,48], cervix[48–50], salivary gland (SG)[51,52], bladder[53–56], trachea[57–59], lacrimal gland (LG)[60–62], esophagus[63–65], prostate[66–69], and urethra[70,71] (Figure 1A).

Among these multilayered epithelia, thickness varies extensively: protective epithelia such as the vagina, esophagus, and cornea form thick, multi-layered barriers, whereas glandular epithelia, including the mammary and salivary glands, are comparatively thin, typically comprising only two cell layers (Figure 1B, C). To correct for batch effects, background noise, and sex-related variation, we applied the Cluster Similarity Spectrum (CSS) integration method[72] (Figure 1D, Figure S1E). The inclusion of three independent datasets per tissue allowed us to confirm that residual batch effects were minimal compared with genuine biological variation (Figure S1F–H). The resulting atlas includes not only epithelial cells but also stromal, immune, endothelial, neural, and erythroid populations (Figure 1E, F). Unsupervised clustering identified the major cell class across tissues, which were further resolved into 25 transcriptionally distinct transcriptomic clusters (Figure 1G; Figure S1I).

Remarkably, epithelial cells exhibited a much higher degree of transcriptional heterogeneity and organ specificity than stromal, endothelial, or immune cells (Figure 1E, G). To quantify this

effect, we computed pairwise the cosine similarity between transcriptomes across all analyzed organs. Hierarchical clustering of these similarity scores confirmed that epithelial cells display the highest level of tissue specificity (Figure 1H).

Analyses of this atlas reveal that epithelial cells exhibit greater transcriptional heterogeneity and organ specificity than stromal or immune populations, establishing a high-resolution foundation to unravel the gene regulatory circuits that shape multilayered epithelia.

Hierarchical basal–suprabasal transcriptional architecture is shared across organs

To resolve epithelial heterogeneity, we focused our analysis on the epithelial compartments of the integrated atlas. Tissue-specific sub clustering and marker-based annotation successfully recovered all known epithelial cell types across the 14 multilayered epithelia (Figure 2A; S2A).

Remarkably, all tissues shared a basal cell (BC) population and one or more suprabasal populations of progenitors and differentiated cells, collectively referred to as suprabasal cells (SBCs).

After integration, epithelial cells failed to cluster by germ layer despite their diverse embryonic origins (Figure 2B, C). Instead, cells better grouped by physiological function and, most prominently, by tissue architecture (Figure 2D, E; Figure S2B). Classification into squamous, reticular, or columnar epithelia revealed the strongest separation in integrated space, quantitatively validated by Local Inverse Simpson's Index (LISI) analysis (Figure 2F; see Methods). We next asked whether cells grouped by identity. In the integrated UMAP, BCs from all tissues coalesced into a continuous transcriptional domain, while SBCs formed discrete clusters defined by their architectural subtype: squamous versus reticular/columnar (Figure 2G, H; S2C-E).

To systematically dissect how cell identity, epithelial architecture, and tissue context shape these transcriptional differences, we decomposed per-gene variance into independent components using a linear contrast-based framework (Figure S2F, G; see Methods). This approach partitioned deviations from the global mean expression across tissues into biologically interpretable sources. It captured the distinct influence of cell identity (basal vs. suprabasal), epithelial architecture (squamous, columnar, reticular), and tissue-specific factors.

We next assigned genes to expression modules based on their dominant source of variance, ensuring that each module captured the main factor shaping its expression pattern (Figure S2H).

To validate that this hierarchical partitioning reflects genuine biological modularity rather than random structure, we performed permutation testing. Cell identity and tissue labels were randomly reassigned across samples 500 times (See Methods), and the proportion of variance (η^2) explained by each factor was recalculated for each permutation. The observed η^2 values consistently surpassed those from the permuted datasets, confirming that the decomposition into *identity*-, *architecture*-, and *tissue-level* components captures statistically robust, non-random biological structure (Figure 2I). Supporting this, intra-module coherence analysis revealed strong correlation among genes within each module, further validating the decomposition's robustness (Figure S2I).

The model yielded biologically interpretable insights. For instance, the transcriptional profile of corneal suprabasal cells could be reconstructed by integrating a *Suprabasal-core* module with *Squamous*, a *Cornea-core*, and *Cornea-specific suprabasal* modules (Figure 2J). This

architecture suggests a Russian-doll-like organization of cell type transcriptomes: core identity modules form the innermost layer, with progressively specialized modules, linked to architecture, tissue, and cell type, forming concentric outer layers (Figure 2K).

To assess how transcriptomic variation differs between epithelial compartments, we applied the variance decomposition framework to individual cells grouped by identity. This analysis revealed a striking asymmetry. In basal cells, variance was predominantly explained by identity-related effects, suggesting that their transcriptional program is relatively stable across tissues. In contrast, suprabasal cells redistributed variance toward tissue- and cell type–specific components, indicating context-dependent diversification (Figure 2L).

Together, these analyses reveal a hierarchical organization of epithelial transcriptomes: a shared program maintains basal cell identity across all tissues, while suprabasal programs progressively diversify in an architecture- and tissue-specific manner.

A shared p63-centered network defines basal identity and supports tissue-specific programs

The strong identity-driven transcriptional similarity of basal cells across tissues prompted us to ask whether a common GRN maintains basal identity in all multilayered epithelia. Analysis of the integrated basal compartments revealed a transcriptionally homogeneous population expressing canonical basal markers (*Krt5*, *Krt14*, *Krt15*) (Figure 3A, B), including a subset of proliferative cells (Figure S3A–B). The fraction of cycling BCs varied across organs, likely reflecting intrinsic differences in epithelial turnover (Figure S3B).

Using our previously described variance decomposition approach (Figure 2K), we identified ~270 genes consistently enriched in basal cells, defining the *Basal-core* module (Figure 3C). Functional enrichment linked this module to cell-matrix adhesion, proliferation, and PI3K and Hippo pathways (Figure 3D), reflecting a coordinated regulatory program that underlies the defining structural and proliferative features of BCs.

Correlation analysis revealed that the *Basal-core* module forms a densely interconnected network of transcription factors and basal markers. *Trp63* and *Snai2* are at the core of the most consistently co-expressed network across tissues (Figure S3C). For example, *Snai2* correlated positively with *Dll1* and *Lamb3*, while *Trp63* correlated with the integrin genes *Itga6* and *Itgb4* (Figure 3E). Notably, p63-deficient mice fail to develop stratified epithelia[15,16], whereas *Snai2*-deficient mice show only premature differentiation[73], highlighting the functional hierarchy within this basal transcriptional backbone.

In addition, analysis of ΔNp63, the predominant *Trp63* isoform present in basal cells, using ChIP-seq data from mouse keratinocytes[74] revealed a significant enrichment of p63 binding sites among *Basal-core* genes compared to the genomic background (Figure 3F; S3D). This included direct binding to *Krt5* and *Itgb4* and to ~25% of the *Basal-core* genes overall (Figure S3E). Consistently, immunostaining for p63 and KRT5 across all 14 epithelia confirmed that the basal program is maintained at the protein level (Figure 3G; Figure S3F).

Having defined the conserved p63-centered Basal-core program, we asked whether basal cells exhibit tissue-specific variations. We further investigated whether these variations correspond to specialized functions. While some genes displayed tissue-restricted expression within basal compartments, these variations did not coalesce into coherent functional modules. Instead, a

single, recurrent and coordinated program emerged; the formation of contractile basal myoepithelial cells.

These cells express smooth-muscle actin (*Acta2*, encoding for αSMA) alongside several contractility-associated genes such as *Myh11*, *Myl*9*, Mylk* and *Tpm2* (Figure 3H, I; S3G). SMA^+ cells are present in the thymus, the mammary, lacrimal and salivary gland (Figure 3J). Analysis of individual datasets revealed distinct patterns. In the mammary gland and thymus, all BCs uniformly expressed *Acta2*. In contrast, salivary and lacrimal glands contained two basal subtypes: a ductal BC population expressing only *Krt5*, and an acinar BC cluster co-expressing *Krt5* and *Acta2* (Figure 3K).

These analyses define a shared, p63-centered GRN that establishes basal epithelial identity across multilayered epithelia. This core network forms a stable regulatory backbone upon which tissue-specific modules, such as the myoepithelial program, can be overlaid to generate functional diversity.

Modular diversification drives organ-specific suprabasal programs and functions

While p63 governs a common basal program, the organization and tissue-specific variation of suprabasal transcriptional programs across organs remain poorly characterized. Integrative analysis of the suprabasal compartment revealed marked heterogeneity among SBC (Figure 4A), most apparent in keratins expression. Distinct combinations of simple (*Krt8/18/19*), mucosal (*Krt4/13/16*), or cornified (*Krt1/10/6a*) keratins contrast sharply with the uniform *Krt5/14* expression typical of basal cells (Figure 3B; 4B). Quantitative comparison of transcriptomic signatures revealed greater cross-tissue divergence among suprabasal than basal cells (Figure 4C), indicating that diversification is concentrated in this compartment.

In the integrated space, cells segregated according to epithelial architecture; squamous, columnar, or reticular, as observed in the whole epithelial atlas (Figure 4D). To formalize this organization, we applied our variance decomposition model, which partitions gene expression into nested components of shared, architectural, and tissue-specific variance (Figure 4E). This analysis revealed a compact *Suprabasal-core* module, approximately 2.5-fold smaller than the *Basal-core*, enriched for adhesion and cytoskeletal genes regulated by Grainyhead-family factors (*Grhl1*) and *Zfp750* (Figure 4F, G; S4A, B).

On top of this shared core, architecture-specific modules encode epithelial type, and include key transcription factors previously implicated in these epithelia: *Irf6*[75] and *Grhl3*[76] in squamous epithelia; *Relb*[77], *Irf8*[78], and *Nfkb2*[79] in reticular epithelium; and *Nr2f2*[80] and *Creb3l2*[81] in columnar epithelia (Figure 4G, H). The hierarchical structure inferred from variance decomposition was independently corroborated by unbiased motif enrichment analysis, which identified the same transcriptional regulators among the top enriched motifs within each module (Figure 4I, S4C).

Yet tissues with a given architecture retained distinctive signatures. In squamous epithelia, SBCs spanned a spectrum from flexible mucosal keratins (*Krt4/13*), supporting adaptable suprabasal layers, to rigid keratins (*Krt1/10)* and cornified envelope proteins *(Ivl, Lor*), forming robust protective barriers (Figure S4D, E). *Tissue-core* modules varied even more, reflecting local regulatory inputs that overlay organ- and cell type–specific effectors on top of shared architectural modules, such as *Nkx2-1* in tracheal SBCs or *Dll4* in the thymic epithelium (Figure S4F). At the periphery of the hierarchy, *Cell type-specific* modules implement organ functions: *caseins* in the mammary gland, *uroplakins* in the bladder, and *mucins* or *secretoglobins* in the salivary and lacrimal glands (Figure 4J).

Within the framework of our variance decomposition, suprabasal transcriptomes mirror the hierarchical “Russian-doll” organization of basal cells. However, they exhibit markedly greater diversification, layering architectural and tissue-specific programs that encode organ-specific epithelial functions. Together with the conserved basal program, this modular organization allows multilayered epithelia to balance structural cohesion with organ-specific functional specialization.

Notch signaling governs epithelial cell fate commitment across multilayered tissues

Having delineated the modular architecture of basal and suprabasal transcriptomes, we next sought to identify the molecular circuitry that connects them. Within the *Basal-core* GRN, we identified some Notch ligands (*Jag2*, *Dll1* and to a lesser extent *Jag1*) consistently co-expressed with *p63* across all multilayered epithelia (Figure 3H). By contrast, the direct Notch target and downstream effector *Hes1* was restricted to cells at the basal–suprabasal interface (Figure 5A). This sender-cell/receiver-cell topology implied that Notch may act as the molecular switch dictating fate commitment, consistent with its classical role as a universal arbiter of cell-fate decisions[82].

To test this model, we genetically forced the expression of constitutive form of the *Notch1* receptor in BCs using *K5-CreERT2:N1ICD-nGFP* mice[83], and compared them to *K5-CreERT2:mTmG* controls[84]. Across all 14 tissues examined, Notch activation induced a basal-to-suprabasal conversion. Genetically traced mutant cells exited the basal layer and acquiring suprabasal or luminal markers, whereas control cells predominantly retained a basal cell identity (Figure 5B, C; S5A, B). The stability of lineage converted cells depended on tissue context: in rapidly renewing epithelia (Figure S3B), SBCs, including *N1ICD*$^{+}$, were shed following upward migration. By contrast, converted luminal cells persisted or expanded in glandular epithelia

(Figure S5C–E). These experiments indicate that Notch is sufficient to trigger suprabasal or luminal cell commitment, although the outcome of the mutant cells depends on the local tissue context and the turnover of SBCs in each tissue.

We next investigated how Notch is wired into the basal network. ΔNp63 ChIP-seq showed p63 occupancy at the *Jag2* promoter (overlapping previously mapped p63 motifs[85–87]) and at the *Hes1* promoter (Figure 5D–E), consistent with previous observations of p63 activating Notch ligand expression in BCs while repressing the downstream effector *Hes1*[19,87]. This dual regulation delineates two transcriptionally distinct compartments: *p63*$^{+}$*Jag2*$^{+}$ basal "signal-senders" and *p63*$^{-}$*Hes1*$^{+}$ suprabasal "signa-receivers" (Figure 5F). These results define a conserved p63–Notch module that links basal identity to suprabasal fate commitment.

Although previously recognized in the epidermis[17,22], we now show that Notch signaling plays a decisive role in fate commitment across multiple homeostatic multilayered epithelia, including the vagina, cervix, urethra, oral mucosa, esophagus, and bladder. Thus, Notch emerges as a universal scaffold that coordinate progenitor maintenance and lineage segregation across multilayered epithelia.

A conserved GRN underlies vertebrate epithelial stratification

Having established that murine multilayered epithelia share a common set of genes for stratification, rather than evolving tissue-specific solutions, we asked whether this core GRN reflects an evolutionarily conserved program. To address this, we integrated tissue morphology, single-cell transcriptomics, and gene conservation data across metazoans (Figure 6A).

Epithelial stratification first arose as a vertebrate innovation, transforming early protective barriers into specialized tissues with distinct physiological functions. While different vertebrate lineages may have followed distinct paths of diversification, here we focus on the mammalian lineage, taking its organs as a reference (Figure 6B). This transition from simple to multilayered epithelia coincided with the expansion of the p53 gene family and the neofunctionalization of p63[29,88], a key regulator of basal identity (Figure 6C).

To identify innovations underlying this shift, we compared epidermal transcriptomes across species from *Drosophila melanogaster* to mammals[89–94] (Figure 6D). Invertebrate epidermis was single-layered, whereas fish and amphibians developed stratified structures expressing simple epithelial keratins in suprabasal cells. Birds and mammals further elaborated this program with multi-step differentiation and cornification (Figure 6D)[95].

Cross-species GRN analysis revealed asymmetric conservation between compartments (Figure 6E, F, S6A). Orthologs of mouse basal genes including *p63*, *Jag2*, and other basal regulators were broadly present and expressed in the basal epidermal cells across vertebrates (~83%) (Figure 6E). In contrast, suprabasal GRNs showed greater lineage specificity (~60–75%), reflecting independent innovations (Figure 6G; S6B). Some suprabasal regulators, such as *Grainyhead* family members, were preserved across lineages, but most effectors diversified with ecosystems: *keratins* in birds, *mucins* in amphibians, and cornification factors in mammals (Figure 6E; S6C). Comparative analysis of murine *Basal-* and *Suprabasal-core* confirmed that basal genes are more evolutionarily constrained, with ortholog recovery decreasing with phylogenetic distance (Figure 6H).

These analyses reveal that the basal compartment of vertebrate epidermis is deeply conserved, maintaining basal identity through a core p63–Notch program, the same axis we observed in

murine multilayered epithelia. In contrast, suprabasal layers have diversified in each lineage, layering tissue- and biotope-specific modules onto this conserved basal scaffold. This modular architecture links a stable basal regulatory axis to lineage-specific effector programs, exemplifying how vertebrate multilayered epithelia combine structural integrity with adaptive functional specialization across organs and species.

Spatial and temporal partitioning of ancestral and novel gene networks drives epithelial diversification

Building on this framework, we asked how evolutionary innovation is incorporated within multilayered epithelia. Although the basal program is ancient and conserved, suprabasal compartments exhibit extensive transcriptional variation across tissues and species. To examine how innovation arose within this conserved framework, we assessed the phylogenetic age of expressed genes across epithelia. Younger genes were consistently enriched in suprabasal compartments, independent of the overall transcriptome age index (TAI)[96] of each tissue (Figure 7A; S7A–B), indicating that suprabasal layers function as evolutionary hotspots where novel modules can be recruited without disrupting core programs.

Analysis of our 14-tissues gene modules confirmed this trend. Recently evolved genes (phylostrata > 8) were enriched in suprabasal modules, up to sixfold in the lacrimal gland (Figure 7B; S7C). Consistent with this, the *Suprabasal-core* itself was consistently enriched (Figure 7C), whereas the Invariant module was depleted and the *Basal-core* showed no bias. Together, these differences reveal a hierarchical organization of evolutionary constraint: basal programs remain anchored in the ancestral GRN, while suprabasal modules are continually remodeled through assimilation of younger genes.

The same principle extended to morphological organization. The *Squamous-core*, representing the earliest stratified architecture, was dominated by conserved genes, whereas *Reticular* and *Columnar* cores, associated with later-evolving secretory, transitional and immunological epithelia, contained more recently evolved genes (Figure 7D). This evolutionary gradient mirrors the sequential emergence of epithelial specializations, with new gene modules progressively incorporated as tissue architectures diversified in the mammalian lineage.

Two representative gene clusters well illustrate this process. The epidermal differentiation complex (EDC), required for barrier formation, and the casein locus, essential milk components, both contribute disproportionately to the elevated TAI of epidermis and mammary gland (Figure 7E). Each represents a recently expanded locus that has become transcriptionally integrated into pre-existing epithelial GRNs. In the skin, EDC genes are targets of the ancestral regulators p63 and Grainyhead-like (*Grhl*) [97–100], indicating that newly acquired enhancers have linked this cluster to the ancient and conserved stratified-epithelium network[101–103]. In the mammary gland, casein genes, which emerged in early mammals, are co-regulated by broadly conserved factors such as the Notch pathway together with mammal-specific hormonal regulators including Stat5[104,105], also reflecting the addition of new regulatory inputs onto an ancestral scaffold. These examples illustrate how recently evolved clusters are functionally assimilated through enhancer-level integration into existing GRNs, expanding tissue-specific repertoires while preserving overall network architecture.

To test the generalization of this principle, we compared six epithelial gene clusters: *keratins*, *desmosomal cadherins*, *EDC*, *secretoglobins*, *caseins*, and *exocrine-gland secreted peptides* (ESPs). Comparative genomics revealed a consistent trend: ancient clusters (keratins, desmosomal cadherins) are broadly expressed across epithelia, consistent with their conserved

structural function, whereas recent clusters (caseins, ESPs) show restricted, lineage-specific expression (Figure 7F–G; S7D–E).

Collectively, these findings reveal that epithelial evolution is governed by a spatial and temporal partition of regulatory innovation. Basal layers preserve tissue integrity through an ancestral GRN of deeply conserved genes, whereas suprabasal layers incorporate newly emerged genes into dynamic, tissue-specific circuits. Within each epithelium, these ancient and recent programs coexist, coupling evolutionary continuity with adaptive change (Figure 7H). This hierarchical organization provides a unified framework for epithelial evolution, showing how tissues acquire new functions while maintaining a robust ancestral core.

Discussion

Organogenesis of multilayered epithelia relies on a conserved transcriptional program that organizes epithelial cells into a characteristic layered architecture across vertebrate organs. The convergence of multilayered epithelia on this shared architecture raises the question of how this organization emerges across germ layers. The transcription factor p63, long known to orchestrate basal identity in ectodermal epithelia[14–16,106,107], is first expressed during surface ectoderm specification[108] and actively suppresses mesodermal and endodermal fates in vertebrate embryos[109–114]. Yet the repeated activation of p63-associated basal program in mesoderm- and endoderm-derivative to initiate tissues such as the esophagus, prostate, and cervix, shows that it can be assembled *de novo*, redeploying an ancestral ectodermal module outside its lineage of origin. These multiple heterotopies[115], potentially triggered by shared stromal or placode-like inductive cues, illustrate how evolutionary innovation can arise through

redeployment and of pre-existing regulatory circuits, a “toolkit” strategy as proposed by Jacob[116].

While p63 provides a conserved entry point for establishing basal identity across germ layers, its functional output diverges across tissues. In rapidly renewing epithelia such as the epidermis, vagina, and esophagus, basal cells act as multipotent progenitors[117–119], whereas in low-turnover glandular tissues they are mostly unipotent[120–124]. Yet, basal cells show limited diversification in gene expression compared with suprabasal compartments, suggesting that their tissue adaptation occurs primarily through modulation of cellular potency. Interestingly, the diversification of suprabasal gene expression may contribute to this process, as suprabasal ligands have been shown to signal to neighboring basal cells to restrict their potential[125].

Suprabasal diversity emerges from the modular combination of Notch signaling with architecture- and tissue-specific transcriptional programs, generating columnar, squamous or reticular cells and their differentiation into secretory, cornified, or immunological units. How these modules are integrated with Notch remains largely unknown. Defining downstream effectors beyond HES1 is challenging, partly because Notch signaling operates as a multiprotein complex whose tissue-specific composition and cofactors dictate its activating or repressive functions[126,127]. Central open questions are whether Notch selectively engages tissue-specific targets, and to what extent local chromatin states shape its accessibility and regulatory output. Resolving these mechanisms will reveal how a single, conserved pathway can simultaneously coordinate fate commitment and functional diversification.

The ability of Notch signaling to integrate with distinct architectural and tissue-specific programs illustrates a broader evo-devo principle: gene regulatory networks act as combinatorial, reusable

building blocks[128]. Yet, despite extensive morphological diversification, the multilayered epithelial regulatory program has remained deeply conserved for over half a billion years. This persistence likely stems from the essential developmental roles of its core factors, most notably p63, whose mutation disrupts limb, skin, and gland formation across vertebrates[14,15,106]. Multilayered organization may have resolved this constraint by spatially partitioning ancestral and derived regulatory modules, with new suprabasal programs built atop a conserved basal scaffold. This compartmentalized evolvability likely reflects dual selective pressures: developmental robustness preserving the ancestral GRN, and ecological adaptation driving diversification.

This study outlines a developmental, spatial, and evolutionary hourglass encoded in tissue architecture. Early epithelial progenitors from diverse germ layers converge on a conserved and developmentally constrained scaffold that dictate basal identity. From this constrained core, suprabasal programs radiate, integrating tissue- and architecture-specific modules shaped by environmental pressures. By unifying developmental and evolutionary perspectives, our work provides a conceptual framework to understand the emergence and diversification of multilayered epithelia.

Resource availability

Lead contact: Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Robin P. Journot (robinjournot5@gmail.com)

Material availability:

This study did not generate new unique reagents.

Data and code availability

No standardized data types were generated in this study. All data supporting the conclusions can be found in the main text or are available upon request from the corresponding authors, Robin Journot (robinjournot5@gmail.com) and Silvia Fre (silvia.fre@curie.fr). The code used for the bioinformatic analysis will be available on Github upon publication. Any additional information required to reanalyze the data reported in this work paper is available from the corresponding authors upon request.

Acknowledgments

We wish to acknowledge all Fre team members, as well as our colleagues Dr. Yohanns Bellaiche, Dr. Edouard Hannezo and Prof. Michel Labouesse, for technical advice and constructive discussions. The authors acknowledge the Cell and Tissue Imaging Platform PICT-IBiSA, member of the national infrastructure France-Bioimaging (https://ror.org/01y7vt929), supported by the French National Research Agency (ANR-24-INBS-0005 FBI BIOGEN) and the Cell and Tissue Imaging (PICT-IBiSA), Institut Curie, member of France-Bioimaging of the Genetics and Developmental Biology Unit (UMR3215/U934), as well as the In Vivo Experimental Facility, particularly Cedrick Pauchard, Mickael Garcia and Celine Daviaud, for help in the maintenance and care of our mouse colony.

This work was funded by, the French National Research Agency (ANR) grant numbers ANR-21-CE13-0047, ANR-22-CE13-0009 and ANR-24-CE13-3529-01, the Medical Research Foundation FRM "FRM Equipes" EQU201903007821, the FSER (Fondation Schlumberger pour l'éducation et la recherche) FSER20200211117, the Association for Research against Cancer (ARC) label ARCPGA2021120004232_4874, the Worldwide Cancer Research Foundation # 24-

0216, the Ligue contre le Cancer label ELF_FRE_26738 and by Labex DEEP ANR-Number 11-LBX-0044.

R.J. was funded by a PhD fellowship from the French minister of research and by the FRM (FDT202404018370).

C.M. was funded by a postdoc fellowship from ARC (ARCPDF12021020003033) and Labex DEEP.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

R.J., C.M. and S.F. conceived and designed the experiments. R.J., C.M., M.H. and E.K. performed all experiments. M.M.F. contributed to experimental work and edited the manuscript. R.J. performed all bioinformatic analysis. R.J. and C.M. curated, and interpreted results and prepared the figures. S.P. and M.S. provided critical remarks on the manuscript. R.J., C.M. and S.F. wrote the manuscript. S.F. provided funding and project administration. R.J. provided supervision. All authors reviewed and approved the manuscript.

Declaration of interests

The authors declare no competing interests.

Methods

Ethics statement

All studies and procedures involving animals were carried out in strict accordance with the recommendations of the European Community (2010/63/UE) for the protection of vertebrate species used for experimental and other scientific purposes. Approval was granted by the Ethics Committee of the Institut Curie CEEA-IC #2021-029 and the French Ministry of Research (reference #34364-202112151422480). We adhere to the internationally established principles of Replacement, Reduction and Refinement (3Rs) according to the Guide for the Care and Use of Laboratory Animals (NRC 2011). Animal husbandry, supply, maintenance and care in the animal facility of the Institut Curie (facility license #C75-05-18) before and during experiments fully met the needs and welfare of the animals. All animals were housed in individually ventilated cages under a 12:12h light/dark cycle, with water and food available ad libitum. All mice were culled by cervical dislocation. Mice were genotyped by PCR on genomic DNA extracted from an earpiece for adult mice or tail tip for embryos. Mouse breeding and husbandry was managed using the mouse colony organization software MiceManager: https://infenx.com/mouse-colony-management-software/.

In vivo lineage tracing

We crossed the *K5-Cre*^{ERT2} line[121] with *mTmG*[84] or *N1ICD-nGFP*[83] for adult *in vivo* lineage tracing experiments. All mice used were in a mixed genetic background. Reporters and mutant alleles were recombined by a single intraperitoneal injection of tamoxifen free base prepared in sunflower oil containing 10% ethanol (0.1 mg per g of mouse body weight), unless indicated otherwise.

Immunostaining

Tissues samples were fixed in 4% PFA (2 hours at RT), washed (1h in PBS at RT) and incubated 48 hours in 30% sucrose at 4°C before embedding in optimal cutting temperature (OCT) compound. 8 µm thick cryosections (15 µm for the mammary gland) were cut using a cryostat (Leica CM1950) and stored at -20°C. For paraffin immunostaining, tissues were fixed overnight at 4°C in 4% PFA, washed 1 h in PBS and incubated in 70% ethanol until embedding. 5 µm thick sections were cut using a rotational microtome (Thermo scientific, Microm HM340E) and stored at RT. Section were deparaffinized, rehydrated and antigen retrieval was achieved by boiling slides in citrate buffer pH=6 for 15 minutes. Paraffin and frozen sections were blocked (PBS, 5% FBS, 1% BSA) for 1 hour and incubated with primary antibodies diluted in blocking buffer overnight in a humidified chamber. After washing, sections were incubated with a solution of secondary antibodies and DAPI (Merck) diluted in PBS for 1 hour. Finally, the sections were mounted on a slide (Aqua-Polymount, Polysciences). A list of all the antibodies used is provided in Table S1.

Image acquisition

Images were acquired using a LSM780, LSM880 or LSM900 inverted laser scanning confocal microscope (Carl Zeiss) equipped with 25x/0.8 OIL LD LCI PL APO or 40x/1.3 OIL DICII PL APO objectives.

Public transcriptomic data collection

Publicly available scRNAseq/ChIPseq datasets of the 14 murine adult tissues, and epidermis of metazoans were retrieved from the GEO database platform: https://www.ncbi.nlm.nih.gov/geo/ and public repository. All selected scRNA seq datasets were generated using the 10X

technology and the expression matrix were retrieved from the processed output of standard pipeline. A list of all the datasets used can be found in Table S2.

Statistics, reproducibility and data availability

Reproducibility codes for the analyses are available online via GitHub, as detailed in the 'Code availability' section. All statistical methods are described in the corresponding sections of the paper. Atlases of individual tissues and integrated dataset will be shared with associated metadata upon publication. Experiments were performed in biological replicates as indicated. A biological replicate corresponds to different animals. Experiments with at least 3 biological replicates were used to calculate the statistical value of each analysis. All graphs are mean + SD. Statistical analysis were performed using the Wilcoxon test or Chi squared test and are stated in figure legend. *: p-value < 0.05, **: p-value < 0.01, ***: p-value <0.001, ****: p-value < 0.0001.

Transcriptomic dataset pre-processing

Individual datasets were pre-processed individually using Seurat[129] to subset viable and good quality cell transcriptome based on the number, individual transcript detected per droplet and percentage of mitochondrial genes. As dataset originated from various sequencing platform and advance in the technology, the QC threshold were adjusted for each dataset. Globally, cells with too low *nFeatures_RNA*, too low or high *nCount_RNA* or percent.mt > 10 were discarded. In each dataset, low-quality genes were filtered out by excluding features detected in fewer than three cells with >1 UMI, cells with fewer than 500 total UMIs were also removed.

Gene expression counts were normalized using the *SCTransform*() function with UMI correction enabled and the top 3000 variable features retained. Principal Component Analysis (PCA) was performed on the SCT assay with 50 dimensions, and elbow plots were used to select relevant components for downstream analysis. UMAP embedding was computed using the accurate number of principal components, followed by neighborhood graph construction

Cell type annotation

To annotate cell types, we first generated atlases of individual tissues through CSS integration (See Data integration section). The markers of each cluster were identified through differential expression and were compared with known markers. The analysis performed in the original publication of the datasets were qualitatively compared with our annotation to homogenize the process. The number of clusters was adjusted based on resolution of the *FindCluster()* function.

Data integration

To construct a unified adult Multilayered epithelium atlas, we integrated annotated single-cell RNA-seq datasets from 14 distinct tissues: bladder, cervix, cornea, epidermis, esophagus, lacrimal gland, mammary gland, oral mucosa, prostate, salivary gland, thymus, trachea, urethra, and vagina. For each tissue, a previously annotated Seurat object was loaded and processed to harmonize assay formats and reduce memory usage.

To avoid barcode collisions during integration, cell names were prefixed with the tissue label using *RenameCells(*. Non-essential layers, including *scale.data* and *data* slots, were removed using regular-expression-based filtering (*removeLayersByPattern()*), and each object was reduced to its essential components using *DietSeurat().*

Datasets were merged iteratively using Seurat's *merge()* function to optimize memory handling. After each merge, the resulting object was joined using JoinLayers() to align assay layers across datasets. The merged object (Sample) thus represented a unified, layer-aware collection of epithelial cells from all tissues.

The integrated object was normalized using *NormalizeData()*, and the top 3,000 variable genes were identified using the variance-stabilizing transformation (*FindVariableFeatures(method = "vst")*).

To improve global alignment and overcome residual batch effects, we applied *Cluster Similarity Spectrum* (CSS) embedding using *cluster_sim_spectrum()* with Spearman correlation as similarity metric and kernel spectrum construction spectrum_type = 'corr_kernel') ("simspec"[72] package). UMAP was recomputed in CSS space using the first 300 CSS components (*RunUMAP(reduction = "css", dims = 1:300)*), and refined clustering was performed using the first 200 CSS dimensions (*FindNeighbors(dims = 1:200), FindClusters(resolution = 0.1)*).

Evaluation of batch effect correction across replicate

To evaluate the effectiveness of batch correction across epithelial biological replicates, we used the Local Inverse Simpson's Index (LISI) as a quantitative metric for local batch mixing. LISI estimates the diversity of batch labels in each cell's neighborhood, based on the inverse Simpson index of the k-nearest neighbor distribution (See section: *Measure of clustering diversity)*. High LISI values indicate that cells from different replicates are well intermixed, while low values (LISI ≈ 1) indicate that cells co-localize primarily with others from the same batch, suggesting residual batch effects.

To avoid conflating batch effects with true biological variation across tissues, LISI was computed within each tissue type. For each epithelial tissue, we subset the integrated Seurat

object to retain only cells belonging to that tissue. From the corrected CSS embedding (first 30 dimensions), we extracted a low-dimensional matrix for LISI analysis. The *compute_lisi()* function from the lisi package[130] was applied to this embedding, using the replicate label to estimate LISI per cell. For each tissue, we reported the mean LISI across all cells, which provides a robust summary of batch mixing quality. To confirm that improvements were due to the integration procedure, we repeated the same computation using the uncorrected PCA embedding, and compared before-versus-after mean LISI per tissue. Only tissues with at least two biological replicates and more than 50 cells were retained for evaluation. This approach allowed us to assess technical batch removal in a biologically meaningful way, preserving within-tissue variability while excluding inter-tissue differences.

Identification of differentially expressed genes and tissue specific markers

Differentially expressed genes (DEGs) were identified using the *FindMarkers()* function in Seurat, applying the Wilcoxon rank-sum test to compare each epithelial tissue against all others. Only genes expressed in at least 10% of cells in either group and with an average log fold change ≥ 0.25 were retained. P-values were adjusted using the Bonferroni method, and genes with an adjusted p-value < 0.01 were considered significant. The top genes were selected for representation as indicated in figure legends. To define tissue-specific genes, we selected DEGs that were significantly upregulated in one tissue compared to all others and showed minimal or no expression in the remaining tissues (expression in < 10% of cells across other tissues).

Pseudo bulk expression and cosine similarity analysis of cell class

To assess transcriptomic similarity between cell class cells across different Multilayered epithelial tissues, we performed a pseudo bulk expression analysis followed by cosine similarity and hierarchical clustering. Pseudo bulk expression matrices were generated for immune cells using the *AverageExpression()* function in Seurat on a down sampled dataset to a maximum of 1,000 cells per tissue–cell type combination. The resulting matrix was converted to a standard numeric matrix, and cosine similarity between all tissue-level immune cell profiles was computed using the *proxy::simil()* function with the "cosine" method.

The cosine similarity matrix was converted into a distance matrix (Euclidean distance of cosine values) and subjected to hierarchical clustering using the *hclust()* function with the "ward.D2" method.

Measure of clustering diversity

Local Inverse Simpson's Index (LISI) is a metric to quantify local diversity in a low-dimensional embedding like UMAP used in single-cell data analysis to assess cell types are mixed.

Local Inverse Simpson's Index (LISI) was computed to quantify the local diversity of cell classes (e.g., embryonic origin or epithelia architecture) in the integrated UMAP space. For each cell, LISI was calculated based on the label distribution among its k nearest neighbors (typically k = 30).

For each cell, LISI was computed as:

$$LISI = \frac{1}{\Sigma\,(p^2)}$$

where p_i is the proportion of neighbors belonging to category i. The score reflects the effective number of distinct categories in the neighborhood:

LISI = 1 indicates that all neighbors belong to a single category (no diversity).

LISI = N (where N is the number of categories) indicates an equal distribution of neighbors across all categories (maximum diversity).

KEGG pathway enrichment

Basal GRN genes were converted from mouse gene symbols to Entrez identifiers using the *bitr()* function from clusterProfiler[131]. KEGG pathway enrichment was then performed with *compareCluster(fun = "enrichKEGG")* using the *Mus musculus* annotation, applying Benjamini–Hochberg correction ($p < 0.05$; $q < 0.1$). Pairwise semantic similarity between enriched terms was computed with *pairwise_termsim()* to allow redundancy assessment.

Linear model of gene expression and module assignment

Preprocessing

Raw counts were aggregated into pseudo bulk profiles for each unique combination of tissue, cell identity, and cell type. Lowly expressed genes (counts per million < 2 in fewer than 5% of samples) and non-informative features (e.g., mitochondrial genes) were excluded. Normalization was performed using the trimmed mean of M-values (TMM) method to correct for differences in library size.

Cell-Means Model

The expression of each gene g in sample s was modeled using a cell-means approach, where the observed expression is represented as the sum of a global mean and a coefficient specific to the combination of Architecture (A), Tissue (T), and Identity (I) for that sample:

$$Y_{\{g,s\}} = \mu_g + \theta_{\{g,A(s),T(s),I(s)\}} + \epsilon_{\{g,s\}}$$

Here, $y_{\{g,s\}}$ is the expression of gene g in sample s, μ_g is the baseline expression level for gene g, $\theta_{\{g,A(s),T(s),I(s)\}}$ is the deviation parameter for the specific combination of Function, Tissue, and Identity, and $\epsilon_{\{g,s\}}$ is the residual error. This model assigns a unique coefficient to each Function × Tissue × Identity combination, capturing the mean expression level for that group.

Contrast Decomposition

To interpret the cell-means coefficients biologically, we decomposed each deviation parameter $\beta_{\{g,f,t,I\}}$ into a sum of contribution parameters using linear contrasts:

$$\theta_{\{g,A,T,I\}} = \alpha_{\{g,I\}} + \beta_{\{g,A\}} + \gamma_{\{g,A,T\}} + \alpha\beta_{\{g,A,I\}} + \alpha\gamma_{\{g,A,T,I\}}$$

where:

- $\alpha_{\{g,I\}}$ global identity effect, capturing basal vs. suprabasal differences across all architectures and tissues
- $\beta_{\{g,A\}}$ architectural effect, reflecting the average influence of each epithelial class (squamous, columnar, reticular)
- $\gamma_{\{g,A,T\}}$ tissue effect nested within architecture**,** representing tissue-specific deviations relative to the mean of their architectural class
- $\alpha\beta_{\{g,A,I\}}$ modulation of identity by architecture, describing how basal–suprabasal differences vary across epithelial architectures
- $\alpha\gamma_{\{g,A,T,I\}}$ modulation of identity by tissue, capturing tissue-specific adjustments of basal–suprabasal differences within each tissue architecture

Each component was isolated by applying orthogonal contrast matrices to the fitted coefficients, allowing independent estimation of variance attributable to each biological factor.

Module Assignment

Genes were assigned to functional modules based on the dominant source of variation in their expression, determined by statistical significance (adjusted p-value) and effect size (partial eta-squared, η_p^2) with:

$$\eta_p^2 = \frac{F \times df_1}{F \times df_1 + df_2}$$

, where F is the F-statistic, anddf_1 and df_2 are the degrees of freedom.

Genes were assigned to modules according to a hierarchical decision tree:

- Identity modules: if $\alpha_{\{g,I\}}$ was significant in ≥75% of tissues (adjusted p < 1e−10) and had the largest η_p^2. Genes were labeled Basal or Suprabasal depending on the direction of the effect.
- Architecture modules: if $\beta_{\{g,A\}}$ was dominant (adjusted p < 1e-6, highest η_p^2).
- Tissue modules: if $\gamma_{\{g,A,T\}}$ was dominant (adjusted p < 1e-6).
- Architecture × Identity modules: if $\alpha\beta_{\{g,A,I\}}$ was significant (p < 1e−4) and dominant.
- Tissue × Identity modules: if $\alpha\gamma_{\{g,A,T,I\}}$ was significant (p < 0.1) and dominant.
- Invariant genes: if no effect passed thresholds.

This hierarchical assignment ensured that each gene was associated with the main factor explaining its variance in expression, providing a biologically interpretable partitioning of transcriptomic regulation across epithelial architectures and identities.

Validation of Model Assumptions

To ensure the validity of the linear model, we verified the following assumptions:

- Normality of Residuals:

The normality of residuals was assessed using Q-Q plots, which showed approximate normality with minor deviations at the tails.

- Homoscedasticity:

Residuals vs. fitted values plots were examined to confirm that the residuals were randomly scattered around zero, with no clear patterns, indicating homoscedasticity.

- Heteroscedasticity:

The relationship between the standard deviation of residuals and the mean fitted values exhibited a funnel shape, consistent with RNA-seq data. This heteroscedasticity was addressed by the precision weights applied during the voom transformation, which stabilizes variance across genes.

Implementation

All analyses were performed in R (v4.2.0) using *edgeR* for normalization, *limma* for voom transformation, model fitting, and empirical Bayes moderation.

Permutation test for variance decomposition

To assess whether the observed variance decomposition reflected genuine biological modularity rather than random structure, we implemented a permutation test on the pseudobulk expression profiles. For each gene, partial η^2 values were estimated for the main effects of Identity (η^2_{identity}), Architecture ($\eta^2_{\text{architecture}}$), Tissue ($\eta^2_{\text{tissue}}$), Architecture × Identity ($\eta^2_{\text{architecture x identity}}$) and Tissue × Identity ($\eta^2_{\text{tissue x identity}}$) using linear modelling with *limma/voom* and custom contrast matrices.

We then computed the mean η² across all genes for each variance component. To generate a null distribution, we randomly permuted Tissue and Identity labels across pseudo bulk samples while keeping the expression matrix unchanged, rebuilt the design matrix, and repeated the variance decomposition. This procedure was iterated 500 times, yielding a distribution of mean η² values expected under random labeling. Permutation p-values were calculated as:

$$p = \frac{1 + count\ (null\ \geq\ observed)}{1 + n_{perm}}$$

Observed values were compared against the null distributions using boxplots with the empirical value indicated as a black point.

Cell-level variance attribution analysis

To extend gene-level variance partitioning to the single-cell scale, we developed an expression-weighted η² scoring approach. For each gene g, variance explained by identity, function, tissue, or interaction terms was estimated as $\eta^2_{g,f}$ using the linear models described previously. To obtain cell-level values, we weighted each gene's η² contribution by its normalized expression $E_{\{i,g\}}$ in cell i, and then computed the weighted mean across all expressed genes:

$$S_{\{i,f\}} = \frac{\sum_{\{g\}} E_{\{i,g\}}\ \cdot\ \eta^2_{\{g,f\}}}{\sum_{\{g\}} E_{\{i,g\}}}$$

where $S_{\{i,f\}}$ is the variance attribution score of cells i for factor f. These yields, for each cell, a profile of scores reflecting the relative influence of identity-, function-, and tissue-level effects on its transcriptome.

Gene co-expression inference

Raw UMI count data were extracted from the RNA assay of the Seurat object using the LayerData() function. To identify genes suitable for co-expression analysis, we computed, for each gene, the average expression across all cells in which the gene was detected (i.e., non-zero counts) as well as the percentage of cells expressing the gene. Genes with an average expression >1 in expressing cells and detected in >10% of cells were retained for downstream analysis.

Co-expression analysis was performed using the CS-CORE algorithm[132] on each tissues, restricted to cells in the resting (non-proliferative) state. For each tissue, the CS-CORE function was run using the filtered gene list as input, and pairwise co-expression estimates and associated p-values were computed. Multiple testing correction was applied using the Benjamini–Hochberg procedure. Adjusted p-values were calculated for all gene pairs, but no filtering threshold was applied at this stage.

Module activity inference by sparse regression.

Raw counts from epithelial atlas were aggregated into pseudobulk profiles per tissue–identity group. Genes with CPM > 2 in at least 5% of groups were retained, normalized by TMM (*edgeR*), and transformed to log2-CPM with precision weights (*voom*, *limma*). Expression values were z-scored per gene across tissues. Gene modules were defined from the global module table and restricted to sets with >5 expressed genes. For each tissue pseudobulk, gene expression was decomposed into weighted module activities using lasso regression with non-negative coefficients (*glmnet*), applied to a design matrix containing positive and negative copies of each module to allow signed contributions. Final weights were collapsed into signed scores, z-scored across tissues, and displayed as a heatmap, with columns split by tissue function and rows split into invariant, functional core, tissue-core, and tissue-identity modules.

GRN representation

To infer transcriptional GRNs across species, we combined co-expression analysis with differential expression data from single-cell RNA-seq of epithelial tissues in human, mouse, chicken, zebrafish, Xenopus, lamprey, and Drosophila. For each species, transcription factors (TFs) were retrieved from species-specific TF reference databases.

GRNs were constructed by filtering a species-specific co-expression matrix (see Gene co-expression inference) to retain only genes significantly differentially expressed genes between basal and suprabasal cells. Co-expression scores were threshold to generate adjacency matrices, and too weak links were removed. Resulting networks were defined as undirected graphs with genes as nodes and significant co-expression links as edges, weighted by co-expression strength. Networks were exported in GraphML format and visualized in *Cytoscape* using the Prefuse Force Directed layout.

ChIPseq data visualization for p63 binding sites

To visualize p63 ChIP-seq signal across specific genomic regions, we used the R packages Gviz, rtracklayer, and GenomicRanges. The processed ChIP-seq signal for p63 (GSM2559737) was obtained in bedGraph format. Genomic regions of interest were defined using *GRanges* objects with *mm10* genome coordinates.

To display gene annotations, we used the *GeneRegionTrack* function from the *Gviz* package, built from the *TxDb.Mmusculus.UCSC.mm10.knownGene* annotation. A DataTrack object was created to plot ChIP-seq signal intensity as a histogram over the selected region.

All tracks, including the gene track, ChIP signal track, genome axis, and ideogram, were plotted using the plotTracks() function. Genomic regions were selected based on loci of interest, such as candidate target genes or regulatory elements.

Quantification of promoter-centered p63 binding

Genome-wide p63 ChIP-seq coverage was computed from bedGraph tracks using rtracklayer and converted to a strand-independent RleList capped at the 99.9th percentile to limit outliers. Gene coordinates were obtained from *TxDb.Mmusculus.UCSC.mm10.knownGene*, and promoter regions were defined as ±1 kb around the transcription start site (TSS). For each promoter, the mean normalized ChIP-seq signal was extracted using the *ScoreMatrixBin()* function from the genomation package (bin.num = 1). Promoters with mean signal above the 95th percentile of all promoters were considered "highly bound." The fraction of highly bound promoters was computed for basal-core and background gene sets, and finally compared statistically using Fisher's exact test.

TSS-centered metagene profiles

To visualize binding profiles, the same coverage data were averaged in 50 bp bins within ±2 kb of each TSS using a custom extraction function. Mean signal and ± SEM were plotted for basal-core and background promoters.

Transcription factor motif enrichment analysis

To identify candidate transcriptional regulators associated with architecture-specific gene modules, we performed TF enrichment analysis using the ChIP-X Enrichment Analysis (ChEA3)

database. For each architectural module (*Squamous-, Reticular-,* and *Columnar-core*), we provided the list of module genes as input. Enrichment was computed against integrated ChIP-seq reference datasets combining ENCODE, ReMap, and literature-curated TF–target interactions. TFs were ranked by their combined enrichment score, which integrates mean rank across multiple evidence types. The top enriched TFs were visualized as ranked enrichment plots, and representative regulators for each architecture were annotated in the figure. Colored annotated TFs correspond to genes that belongs to the module.

Evolutionary Age Profiling and Tissue Transcriptome Analysis

Gene orthology and evolutionary age were obtained from the *phylomapr* package, which provides a Phylostratum annotation for Mus musculus genes from Barrera-Redondo et al. 2023[133]. UniProt identifiers were mapped to Ensembl gene IDs and gene symbols using the *biomaRt*[134] package.

Transcriptome Age Index (TAI) Calculation Across Tissues

To assess the evolutionary age of gene expression profiles in adult epithelia, we computed the Transcriptome Age Index (TAI) per tissue. Raw gene expression counts were extracted from the Seurat object, and pseudobulk transcriptomes were generated by summing raw counts across all cells from each tissue. Normalization was performed using *log1p* transformation.

Genes were intersected with the phylogenetic map to retain those with known phylostrata. A *PhyloExpressionSet* was then constructed with rows representing genes and columns representing phylostratum and normalized expression. The TAI was calculated for each tissue using the *TAI()* function from the *myTAI* package[96]. The index reflects the average evolutionary age of expressed genes, weighted by their expression level.

Identification of Evolutionarily Biased Contributors

To identify genes contributing disproportionately to the TAI in each tissue, we computed a contribution score as the product of gene expression and phylostratum value (i.e., older genes contribute less).

Differential Expression and TAI in Cell Types

To compare evolutionary biases between basal and suprabasal compartments, we performed tissue-wise differential expression (DE) analysis. For each tissue containing both basal and suprabasal populations, marker gene detection was carried out with *FindMarkers()* using thresholds of logFC > 0.5 and minimum detection fraction > 0.1. The top 500 DE genes upregulated in each identity (basal or suprabasal) were retained.

For each tissue-compartment pair, pseudobulk transcriptomes were generated by summing raw gene counts across all cells, and normalized by dividing by the number of contributing cells to control for sample size. Genes were annotated with their phylostratum and the TAI was calculated as previously described. Briefly, TAI was defined as the weighted mean phylostratum of expressed genes, where gene weights correspond to normalized expression levels. Samples with fewer than 30 annotated genes were excluded.

To account for sampling variability, we applied a bootstrap strategy. For each pseudobulk sample, the set of DE genes was resampled with replacement 150 times, and TAI was recomputed in each replicate. This produced a bootstrap distribution of TAI values per tissue and identity. From these distributions, we estimated mean TAI, standard deviation, and 95% confidence intervals. Differences between basal and suprabasal bootstrap distributions were assessed using Wilcoxon rank-sum tests. Multiple testing correction across tissues was performed using the Benjamini–Hochberg method.

Venn Diagram Analysis

For exploratory purposes, Venn diagrams were generated to visualize overlap in top evolutionary contributors across tissues using *nVennR*. Gene sets were grouped by tissue and displayed interactively, with gene region intersections extracted for interpretation.

Phylostratum enrichment of gene modules

To assess whether specific gene modules were enriched in ancient or recently evolved genes, we annotated each gene with its phylostratum. Genes were binned into three evolutionary age classes: Very Ancient (Phylostratum ≤2), Intermediate (3≤ Phylostratum ≤8), and Recent (Phylostratum >8). Each module, defined as the set of genes assigned to a given identity/function/tissue component, was intersected with these age classes.

For each module and age class, enrichment was evaluated relative to a background set of expressed genes matched to the same tissue and identity context. Genes were considered expressed if they were detected in at least 25% of cells from the corresponding. Enrichment was tested by constructing 2×2 contingency tables of module vs. background membership and applying one-sided Fisher's exact tests. To assess robustness, module fractions were compared against bootstrap resampling distributions (n=100) drawn from the same gene set, and significance was additionally tested with one-sample Wilcoxon tests against the background fraction.

Supplementary Information

Table S1: Antibodies

Antibody	Source	Reference	Dilution
Rabbit anti-K5	Covance	Cat# PRB-160P-100; RRID: AB_291581	1:300

Chicken anti-K5	Biolegend	#905901	1:200
Rat anti-K8	Developmental Studies Hybridoma Bank, University of Iowa	Cat# TROMA-I: RRID: AB_531826	1:300
Mouse anti-p63	Abcam	Cat# ab735; RRID: AB_305870	1:200
Rabbit anti-Ki67	Thermo Fisher Scientific	Cat# MA5-14520; RRID:AB_10979488	1:300
Rabbit anti-HES1	Cell Signaling technology	#11988	1:100
Rabbit anti-GFP		Recombinant anti GFP antibody	1:200
Human anti-GFP		Recombinant anti GFP antibody	1:200
Guinea Pig anti-K13	Progen	Cat# GP-K13; RRID: AB_1541010	1:300
Guinea Pig anti-K10	Progen	Cat# GP-K10; RRID: AB_2920688	1:300
Rabbit anti-LORICRIN	Covance	Cat#: PRB-145P; RRID: AB_10064155	1:400
Rabbit anti-SMA	Abcam	Cat# ab5694; RRID: AB_2223021	1:300
Goat anti-rabbit AlexaFluor-coupled to different fluorochromes (Cy3, Cy5, A488, A405)	Invitrogen-Thermo Fisher Scientific	Cat# A10520; RRID: AB_2534029 Cat#A10523; RRID: AB_2534032, Cat# A-11034; RRID: AB_2576217 Cat# A-31556; RRID: AB_221605	1:500
Goat anti-rat AlexaFluor-coupled to different fluorochromes (Cy3, Cy5, A488)	Invitrogen-Thermo Fisher Scientific	Cat# A10522; RRID: AB_2534031, Cat#A10525; RRID: AB_2534034, Cat# A-11006; RRID: AB_2534074	1:500
Goat anti-mouse AlexaFluor-coupled to different fluorochromes (Cy3, Cy5, A488)	Invitrogen-Thermo Fisher Scientific	Cat# A10521; RRID: AB_2534030, Cat# A10524; RRID: AB_2534033 Cat# A-11001; RRID: AB_2534069	1:500
Goat anti-guinea pig AlexaFluor-coupled to different fluorochromes (A568)	Invitrogen-Thermo Fisher Scientific	CAT#A-11075; RRID: AB_2337423	1:500

Table S2: Public transcriptomic dataset

Repository	Tissue	Reference	Stage (weeks)	Datasets
GSE129845	Bladder	Yu et al. 2019 [53]	10	GSM3723360, GSM3723361
GSE180128	Bladder	Baker et al. 2021 [54]	88,12,48,48,96	GSM5453638, GSM5453639, GSM5453640, GSM5453641, GSM5453642
GSE141348	Bladder	Kim et al. 2020 [55]	9	GSM4201633
GSE163029	Bladder	Li et al. 2021 [56]	9	GSM4970435
GSE269056	Lacrimal gland	Huang et al. 2024 [60]	13	GSM8305673, GSM8305674,

				GSM8305675, GSM8305676, GSM8305677, GSM8305678
GSE232146	Lacrimal gland	Delcroix et al. 2023 [61]	8	GSM7316423, GSM7316424, GSM7316425
GSE132420	Lacrimal gland	Rattner et al. 2023 [62]	16	GSM3863114, GSM3863116, GSM3863118, GSM3863120, GSM3863122, GSM3863123, GSM3863124, GSM3863125
GSE167992	Cornea	Altschuler et al. 2021 [34]	11	GSM5121234
GSE263994	Cornea	Jiang et al. 2024 [35]	7	GSM8208504, GSM8208505, GSM8208506
GSE272702	Cornea	Efraim et al. 2024 [36]	7	GSM8409444
GSE142471	Epidermis	Haensel et al. 2020 [31]	7	GSM4230076, GSM4230077
GSE129218	Epidermis	Joost et al. 2020 [32]	9	GSM4186891, GSM4186892, GSM4186893
GSE201447	Epidermis	Paldor et al. 2022 [33]	8	GSM6064614
GSE220065	Esophagus	Vercauteren Drubbel and Beck. 2023 [63]	8	GSM6784567, GSM6784568, GSM6784569
GSE193376	Esophagus	Kabir et al. 2022 [64]	12	GSM5799214, GSM5799218, GSM5799222, GSM5799226
E-MTAB-8662	Esophagus	McGinn et al. 2021 [65]	4,4,4,10,10,10	P28_1, p28_2, p28_3, Adult_1, Adult_2, Adult_3
GSE106273	Mammary gland	Bach et al. 2017 [43]	8	GSM2834498, GSM2834499
GSE150580	Mammary gland	Li et al. 2020 [44]	12	GSM4552664, GSM4552665, GSM4552666
GSE164017	Mammary gland	Cheng et al. 2023 [45]	2.5, 5, 5	GSM4994963, GSM4994964, GSM4994965
GSE191219	Mammary gland	Saeki et al. 2024 [46]	43	GSM5741807
GSE120654	Oral mucosa	Jones et al. 2019 [37]	12	GSM3407039, GSM3407040
GSE228635	Oral mucosa	Kondo et al. 2023 [38]	10	GSM7134242
GSE269575	Oral mucosa	Jaber et al. 2024 [39]	8	GSM8322397, GSM8322398
GSE146811	Prostate	Karthaus et al. 2020 [66]	10	GSM4474184, GSM4474185, GSM4474186, GSM4474187, GSM4474188, GSM4474189, GSM4474190
GSE150692	Prostate	Crowley et al. 2020 [67]	9	GSM4556594, GSM4556595, GSM4556596,

				GSM4556597, GSM4556598, GSM4556599
GSE165741	Prostate	Graham et al. 2023 [69]	24	GSM5049201, GSM5049204, GSM5049207, GSM5049210
GSE224452	Prostate	Aparicio et al. 2025 [68]	10	GSM7024431, GSM7024432
GSE150327	Salivary gland	Hauser et al. 2020 [51]	4, 4, 70	GSM4546896, GSM4546897, GSM4546898
GSE175649	Salivary gland	Horeth et al. 2021 [52]	49	GSM5343233
GSE131339	Thymus	Cowan et al. 2019 [40]	8	GSM3770663, GSM3770664, GSM3770669, GSM3770670
GSE228198	Thymus	Horie et al. 2023 [42]	9	GSM7116509
GSE240016	Thymus	Kousa et al. 2024 [41]	8, 8, 8, 8, 72, 72, 72, 72	GSM7679869, GSM7679870, GSM7679871, GSM7679872, GSM7679885, GSM7679886, GSM7679887, GSM7679888
GSE240020	Thymus	Kousa et al. 2024 [41]	80	GSM7963648
GSE103354	Trachea	Montoro et al. 2018 [57]	9	GSM3314348, GSM3314349, GSM3314350, GSM3314351, GSM3314352, GSM3314353
GSE152692	Trachea	Kiyokawa et al. 2021 [58]	8	GSM4623795
GSE244215	Trachea	Lin et al. 2024 [59]	8, 96, 8, 96	GSM7808936, GSM7808937, GSM7808938, GSM7808939
GSE145865	Urethra	Joseph et al. 2020 [70]	10	GSM4338168, GSM4338169
GSE173096	Urethra	Joseph et al. 2021 [71]	10	GSM5258273, GSM5258274, GSM5258275
GSE201274	Vagina	Kang et al. 2023 [47]	15	GSM6057164
E-MTAB-11491	Vagina	Winkler et al. 2024 [48]	12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 72, 72	Ind001_vagina01, Ind002_vagina02, Ind003_vagina03, Ind007_vagina07, Ind007_vagina08, Ind007_vagina09, Ind007_vagina10, Ind007_vagina11, Ind007_vagina12, Ind007_vagina13, Ind007_vagina14, Ind007_vagina15, Ind001_vagina01, Ind001_vagina01
E-MTAB-11491	Cervix	Winkler et al. 2024 [48]	12	Ind001-cervix1, Ind001-cervix2,

				Ind001-cervix3, Ind001-cervix4, Ind001-cervix5, Ind001-cervix6, Ind001-cervix7, Ind001-cervix8, Ind001-cervix9, Ind001-cervix10
GSE128987	Cervix	Chumduri et al. 2021 [49]	15	GSM3689562
GSE196529	Cervix	Cooley et al. 2023 [50]	16	GSM5888614
GSE97213	Epidermis, ChIP seq p63	Fan et al. 2018[74]	New born	GSM2559737
	Epidermis, Drosopjila melanogaster	Li et al. 2022[89]	Adult	body_wall_raw
	Epidermis, Lamprey	Pang et al. 2025[92]	Larvae	Epidermis
GSE224695	Epidermis, Danio rerio	Aman et al.[94]	3 weeks PF	GSM7029635
GSE195790	Epidermis, Xenopus laevis	Liao et al. 2022[93]	Adult	GSM6214285
GSE179690	Epidermis, Gallus gallus	Lachner et al. 2022[90]	New born	GSM5428157
GSE162185	Epidermis, Homo sapiens	Gao et al. 2021[91]	Adult	GSM4946163

Resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-K5	Covance	Cat# PRB-160P-100; RRID: AB_291581
Chicken anti-K5	Biolegend	#905901
Rat anti-K8	Developmental Studies Hybridoma Bank, University of Iowa	Cat# TROMA-I: RRID: AB_531826
Mouse anti-p63	Abcam	Cat# ab735; RRID: AB_305870
Chicken anti-K14	Biolegend	# 906004
Rabbit anti-Ki67	Invitrogen-Thermo Fisher Scientific	Cat# MA5-14520; RRID:AB_10979488
Rabbit anti-HES1	Cell Signaling technology	#11988
Human anti-GFP	Institut Curie antibodies platform	NA
Rabbit anti-GFP	Institut Curie antibodies platform	NA
Guinea Pig anti-K13	Progen	Cat# GP-K13; RRID: AB_1541010
Guinea Pig anti-K10	Progen	Cat# GP-K10; RRID: AB_2920688

Rabbit anti-LORICRIN	Covance	Cat#: PRB-145P; RRID: AB_10064155
Goat anti-rabbit AlexaFluor-coupled to different fluorochromes Cy5	Invitrogen-Thermo Fisher Scientific	Cat#A10523; RRID: AB_2534032
Goat anti-rabbit AlexaFluor-coupled to different fluorochromes A488	Invitrogen-Thermo Fisher Scientific	Cat# A-11034; RRID: AB_2576217
Goat anti-rabbit AlexaFluor-coupled to different fluorochromes A405	Invitrogen-Thermo Fisher Scientific	Cat# A-31556; RRID: AB_221605
Goat anti-mouse AlexaFluor-coupled to different fluorochromes Cy3	Invitrogen-Thermo Fisher Scientific	Cat# A10521; RRID: AB_2534030
Goat anti-mouse AlexaFluor-coupled to different fluorochromes Cy5	Invitrogen-Thermo Fisher Scientific	Cat# A10524; RRID: AB_2534033
Goat anti-mouse AlexaFluor-coupled to different fluorochromes A488	Invitrogen-Thermo Fisher Scientific	Cat# A-11001; RRID: AB_2534069
Goat anti-rat AlexaFluor-coupled to different fluorochromes Cy3	Invitrogen-Thermo Fisher Scientific	Cat# A10522; RRID: AB_1500704 ,
Goat anti-rat AlexaFluor-coupled to different fluorochromes Cy5	Invitrogen-Thermo Fisher Scientific	Cat# A1052; RRID: AB_10562712
Goat anti-rat AlexaFluor-coupled to different fluorochromes A488	Invitrogen-Thermo Fisher Scientific	Cat# A-11006; RRID: AB_141373
Goat anti-guinea pig AlexaFluor-coupled to different fluorochromes (A568)	Invitrogen-Thermo Fisher Scientific	CAT#A-11075; RRID: AB_2337423
Chemicals, peptides, and recombinant proteins		
Sunflower Oil	Merck-Millipore	S5007
Tamoxifen	MP Biomedicals	0215673891
Paraformaldehyde (PFA)	Electron Microscopy Sciences	15710
Aqua poly/mount	Tebu Bio	07918606-5
Glycerol	Euromedex	15710
Triton X-100	Euromedex	2000C
Sucrose	Sigma	S0389
Tissue-Tek O.C.T.	Sakura	4583
Experimental models: Organisms/strains		
K5-CreERT2	Van Keymeulen et al.[8]	RRID:IMSR_JAX:029155,
R26-N1ICD	Murtaugh et t al.[81]	RRID:IMSR_JAX:008159
R26-mTmG	Muzumdar et al.[71]	RRID:IMSR_JAX:037456
Software and algorithms		
R v4.0.3	R Core Team	https://www.r-project.org/
Python v3.7.7	Python Software Foundation	https://python.org/
Seurat v4.2.0	Satija Lab	https://satijalab.org/seurat/
ImageJ	National Institutes of Health	https://imagej.nih.gov/ij/

Zen blue 3.5 (ZEISS ZEN lite)	Carl Zeiss Microscopy	https://www.zeiss.com/microscopy/en/products/software/zeiss-zen-lite.html

Reference

1. Simpson, C.L., Patel, D.M., and Green, K.J. (2011). Deconstructing the skin: cytoarchitectural determinants of epidermal morphogenesis. Nat. Rev. Mol. Cell Biol. *12*, 565–580. https://doi.org/10.1038/nrm3175.

2. Gjorevski, N., and Nelson, C.M. (2011). Integrated morphodynamic signalling of the mammary gland. Nat. Rev. Mol. Cell Biol. *12*, 581–593. https://doi.org/10.1038/nrm3168.

3. Kadouri, N., Nevo, S., Goldfarb, Y., and Abramson, J. (2020). Thymic epithelial cell heterogeneity: TEC by TEC. Nat. Rev. Immunol. *20*, 239–253. https://doi.org/10.1038/s41577-019-0238-0.

4. Ono-Minagi, H., Nohno, T., Serizawa, T., Usami, Y., Sakai, T., Okano, H., and Ohuchi, H. (2023). The Germinal Origin of Salivary and Lacrimal Glands and the Contributions of Neural Crest Cell-Derived Epithelium to Tissue Regeneration. Int. J. Mol. Sci. *24*, 13692. https://doi.org/10.3390/ijms241813692.

5. Creuzet, S., Vincent, C., and Couly, G. (2005). Neural crest derivatives in ocular and periocular structures. Int. J. Dev. Biol. *49*, 161–171. https://doi.org/10.1387/ijdb.041937sc.

6. Hughes, M.W., and Chuong, C.-M. (2003). A Mouthful of Epithelial–Mesenchymal Interactions. J. Invest. Dermatol. *121*, vii–viii. https://doi.org/10.1111/j.1523-1747.2003.12651.x.

7. Roly, Z.Y., Backhouse, B., Cutting, A., Tan, T.Y., Sinclair, A.H., Ayers, K.L., Major, A.T., and Smith, C.A. (2018). The cell biology and molecular genetics of Müllerian duct development. WIREs Dev. Biol. *7*, e310. https://doi.org/10.1002/wdev.310.

8. Kurita, T. (2010). Developmental Origin of Vaginal Epithelium. Differ. Res. Biol. Divers. *80*, 99–105. https://doi.org/10.1016/j.diff.2010.06.007.

9. Lungova, V., Verheyden, J.M., Herriges, J., Sun, X., and Thibeault, S.L. (2015). Ontogeny of the mouse vocal fold epithelium. Dev. Biol. *399*, 263–282. https://doi.org/10.1016/j.ydbio.2014.12.037.

10. Figueiredo, M., Zilhão, R., and Neves, H. (2020). Thymus Inception: Molecular Network in the Early Stages of Thymus Organogenesis. Int. J. Mol. Sci. *21*, 5765. https://doi.org/10.3390/ijms21165765.

11. Kwon, O.-J., and Xin, L. (2014). Prostate epithelial stem and progenitor cells. Am. J. Clin. Exp. Urol. *2*, 209–218.

12. Melino, G., Memmi, E.M., Pelicci, P.G., and Bernassola, F. (2015). Maintaining epithelial stemness with p63. Sci. Signal. *8*, re9–re9. https://doi.org/10.1126/scisignal.aaa1033.

13. Senoo, M., Pinto, F., Crum, C.P., and McKeon, F. (2007). p63 Is Essential for the Proliferative Potential of Stem Cells in Stratified Epithelia. Cell *129*, 523–536. https://doi.org/10.1016/j.cell.2007.02.045.

14. Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R.T., Tabin, C., Sharpe, A., Caput, D., Crum, C., et al. (1999). p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. Nature *398*, 714–718. https://doi.org/10.1038/19539.

15. Mills, A.A., Zheng, B., Wang, X.-J., Vogel, H., Roop, D.R., and Bradley, A. (1999). p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature *398*, 708–713. https://doi.org/10.1038/19531.

16. Koster, M.I., Kim, S., Mills, A.A., DeMayo, F.J., and Roop, D.R. (2004). p63 is the molecular switch for initiation of an epithelial stratification program. Genes Dev. *18*, 126–131. https://doi.org/10.1101/gad.1165104.

17. Blanpain, C., Lowry, W.E., Pasolli, H.A., and Fuchs, E. (2006). Canonical notch signaling functions as a commitment switch in the epidermal lineage. Genes Dev. *20*, 3022–3035. https://doi.org/10.1101/gad.1477606.

18. Cockburn, K., Annusver, K., Gonzalez, D.G., Ganesan, S., May, D.P., Mesa, K.R., Kawaguchi, K., Kasper, M., and Greco, V. (2022). Gradual differentiation uncoupled from cell cycle exit generates heterogeneity in the epidermal stem cell layer. Nat. Cell Biol. *24*, 1692–1700. https://doi.org/10.1038/s41556-022-01021-8.

19. Nguyen, B.-C., Lefort, K., Mandinova, A., Antonini, D., Devgan, V., Gatta, G.D., Koster, M.I., Zhang, Z., Wang, J., Vignano, A.T. di, et al. (2006). Cross-regulation between Notch and p63 in keratinocyte commitment to differentiation. Genes Dev. *20*, 1028–1042. https://doi.org/10.1101/gad.1406006.

20. Watt, F.M., Estrach, S., and Ambler, C.A. (2008). Epidermal Notch signalling: differentiation, cancer and adhesion. Curr. Opin. Cell Biol. *20*, 171–179. https://doi.org/10.1016/j.ceb.2008.01.010.

21. Williams, S.E., Beronja, S., Pasolli, H.A., and Fuchs, E. (2011). Asymmetric Cell Divisions Promote Notch-Dependent Epidermal Differentiation. Nature *470*, 353–358. https://doi.org/10.1038/nature09793.

22. Estrach, S., Cordes, R., Hozumi, K., Gossler, A., and Watt, F.M. (2008). Role of the Notch Ligand Delta1 in Embryonic and Adult Mouse Epidermis. J. Invest. Dermatol. *128*, 825–832. https://doi.org/10.1038/sj.jid.5701113.

23. Lilja, A.M., Rodilla, V., Huyghe, M., Hannezo, E., Landragin, C., Renaud, O., Leroy, O., Rulands, S., Simons, B.D., and Fre, S. (2018). Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. Nat. Cell Biol. *20*, 677–687. https://doi.org/10.1038/s41556-018-0108-1.

24. Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinarsky, V., Cho, J.L., Breton, S., Sahay, A., et al. (2013). Dedifferentiation of committed epithelial cells into stem cells in vivo. Nature *503*, 218–223. https://doi.org/10.1038/nature12777.

25. Van Batavia, J., Yamany, T., Molotkov, A., Dan, H., Mansukhani, M., Batourina, E., Schneider, K., Oyon, D., Dunlop, M., Wu, X.-R., et al. (2014). Bladder cancers arise from distinct urothelial sub-populations. Nat. Cell Biol. *16*, 982–991. https://doi.org/10.1038/ncb3038.

26. Journot, R.P., Huyghe, M., Barthelemy, A., Couto-Moreira, H., Deshayes, T., Harari, L., Sumbal, J., Faraldo, M.M., Dubail, M., Fouillade, C., et al. (2025). Conserved signals control self-organization and symmetry breaking of murine bilayered epithelia during development and regeneration. Dev. Cell. https://doi.org/10.1016/j.devcel.2025.06.007.

27. Merle, C., Rodrigues, C., Pourkhalili Langeroudi, A., Journot, R., Rost, F., Dang, Y., Rulands, S., and Fre, S. (2025). Transcriptional landscapes underlying Notch-induced lineage conversion and plasticity of mammary basal cells. EMBO J. https://doi.org/10.1038/s44318-025-00424-1.

28. Alaminos, M., and Campos, A. (2023). The origin of human epithelial tissue. Histol. Histopathol. *38*, 127–138. https://doi.org/10.14670/HH-18-485.

29. Belyi, V.A., and Levine, A.J. (2009). One billion years of p53/p63/p73 evolution. Proc. Natl. Acad. Sci. *106*, 17609–17610. https://doi.org/10.1073/pnas.0910634106.

30. Vanbokhoven, H., Melino, G., Candi, E., and Declercq, W. (2011). p63, a Story of Mice and Men. J. Invest. Dermatol. *131*, 1196–1207. https://doi.org/10.1038/jid.2011.84.

31. Haensel, D., Jin, S., Sun, P., Cinco, R., Dragan, M., Nguyen, Q., Cang, Z., Gong, Y., Vu, R., MacLean, A.L., et al. (2020). Defining Epidermal Basal Cell States during Skin Homeostasis and Wound Healing Using Single-Cell Transcriptomics. Cell Rep. *30*, 3932-3947.e6. https://doi.org/10.1016/j.celrep.2020.02.091.

32. Joost, S., Annusver, K., Jacob, T., Sun, X., Dalessandri, T., Sivan, U., Sequeira, I., Sandberg, R., and Kasper, M. (2020). The Molecular Anatomy of Mouse Skin during Hair Growth and Rest. Cell Stem Cell *26*, 441-457.e7. https://doi.org/10.1016/j.stem.2020.01.012.

33. Paldor, M., Levkovitch-Siany, O., Eidelshtein, D., Adar, R., Enk, C.D., Marmary, Y., Elgavish, S., Nevo, Y., Benyamini, H., Plaschkes, I., et al. (2022). Single-cell transcriptomics reveals a senescence-associated IL-6/CCR6 axis driving radiodermatitis. EMBO Mol. Med. *14*, e15653. https://doi.org/10.15252/emmm.202115653.

34. Altshuler, A., Amitai-Lange, A., Tarazi, N., Dey, S., Strinkovsky, L., Hadad-Porat, S., Bhattacharya, S., Nasser, W., Imeri, J., Ben-David, G., et al. (2021). Discrete limbal epithelial stem cell populations mediate corneal homeostasis and wound healing. Cell Stem Cell *28*, 1248-1261.e8. https://doi.org/10.1016/j.stem.2021.04.003.

35. Jiang, H., Liu, M., Yang, W., Hong, Y.-K., Xu, D., Nalbant, E.K., Clutter, E.D., Foroozandeh, P., Kaplan, N., Wysocki, J., et al. (2024). Activation of limbal epithelial proliferation is partly controlled by the ACE2-LCN2 pathway. iScience *27*, 110534. https://doi.org/10.1016/j.isci.2024.110534.

36. Efraim, Y., Chen, F.Y.T., Niknezhad, S.V., Pham, D., Cheong, K.N., An, L., Sinada, H., McNamara, N.A., and Knox, S.M. (2024). Rebuilding the autoimmune-damaged corneal stroma through topical lubrication. BioRxiv Prepr. Serv. Biol., 2024.11.29.626078. https://doi.org/10.1101/2024.11.29.626078.

37. Jones, K.B., Furukawa, S., Marangoni, P., Ma, H., Pinkard, H., D'Urso, R., Zilionis, R., Klein, A.M., and Klein, O.D. (2019). Quantitative Clonal Analysis and Single-Cell Transcriptomics Reveal Division Kinetics, Hierarchy, and Fate of Oral Epithelial Progenitor Cells. Cell Stem Cell *24*, 183-192.e8. https://doi.org/10.1016/j.stem.2018.10.015.

38. Kondo, T., Gleason, A., Okawa, H., Hokugo, A., and Nishimura, I. (2023). Mouse gingival single-cell transcriptomic atlas identified a novel fibroblast subpopulation activated to guide oral barrier immunity in periodontitis. eLife *12*, RP88183. https://doi.org/10.7554/eLife.88183.

39. Jaber, Y., Sarusi-Portuguez, A., Netanely, Y., Naamneh, R., Yacoub, S., Saar, O., Darawshi, N., Eli-Berchoer, L., Shapiro, H., Elinav, E., et al. (2024). Gingival spatial analysis reveals geographic immunological variation in a microbiota-dependent and -independent manner. NPJ Biofilms Microbiomes *10*, 142. https://doi.org/10.1038/s41522-024-00625-2.

40. Cowan, J.E., Malin, J., Zhao, Y., Seedhom, M.O., Harly, C., Ohigashi, I., Kelly, M., Takahama, Y., Yewdell, J.W., Cam, M., et al. (2019). Myc controls a distinct transcriptional program in fetal thymic epithelial cells that determines thymus growth. Nat. Commun. *10*, 5498. https://doi.org/10.1038/s41467-019-13465-y.

41. Kousa, A.I., Jahn, L., Zhao, K., Flores, A.E., Acenas, D., Lederer, E., Argyropoulos, K.V., Lemarquis, A.L., Granadier, D., Cooper, K., et al. (2024). Age-related epithelial defects limit

thymic function and regeneration. Nat. Immunol. *25*, 1593–1606. https://doi.org/10.1038/s41590-024-01915-9.

42. Horie, K., Namiki, K., Kinoshita, K., Miyauchi, M., Ishikawa, T., Hayama, M., Maruyama, Y., Hagiwara, N., Miyao, T., Murata, S., et al. (2023). Acute irradiation causes a long-term disturbance in the heterogeneity and gene expression profile of medullary thymic epithelial cells. Front. Immunol. *14*, 1186154. https://doi.org/10.3389/fimmu.2023.1186154.

43. Bach, K., Pensa, S., Grzelak, M., Hadfield, J., Adams, D.J., Marioni, J.C., and Khaled, W.T. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat. Commun. *8*, 2128. https://doi.org/10.1038/s41467-017-02001-5.

44. Li, C.M.-C., Shapiro, H., Tsiobikas, C., Selfors, L.M., Chen, H., Rosenbluth, J., Moore, K., Gupta, K.P., Gray, G.K., Oren, Y., et al. (2020). Aging-Associated Alterations in Mammary Epithelia and Stroma Revealed by Single-Cell RNA Sequencing. Cell Rep. *33*, 108566. https://doi.org/10.1016/j.celrep.2020.108566.

45. Cheng, J., Smyth, G.K., and Chen, Y. (2023). Unraveling the timeline of gene expression: A pseudotemporal trajectory analysis of single-cell RNA sequencing data. F1000Research *12*, 684. https://doi.org/10.12688/f1000research.134078.2.

46. Saeki, K., Ha, D., Chang, G., Mori, H., Yoshitake, R., Wu, X., Wang, J., Wang, Y.-Z., Wang, X., Tzeng, T., et al. (2024). Perimenopausal and Menopausal Mammary Glands In A 4-Vinylcyclohexene Diepoxide Mouse Model. J. Mammary Gland Biol. Neoplasia *29*, 15. https://doi.org/10.1007/s10911-024-09569-x.

47. Kang, I., Park, J.H., Kim, Y., Kwon, M.S., and Lee, H.K. (2023). Protocol for analyzing γδ T cells from the vagina of diet-induced obese mice using single-cell RNA sequencing and flow cytometry. STAR Protoc. *4*, 102278. https://doi.org/10.1016/j.xpro.2023.102278.

48. Winkler, I., Tolkachov, A., Lammers, F., Lacour, P., Daugelaite, K., Schneider, N., Koch, M.-L., Panten, J., Grünschläger, F., Poth, T., et al. (2024). The cycling and aging mouse female reproductive tract at single-cell resolution. Cell *187*, 981-998.e25. https://doi.org/10.1016/j.cell.2024.01.021.

49. Chumduri, C., Gurumurthy, R.K., Berger, H., Dietrich, O., Kumar, N., Koster, S., Brinkmann, V., Hoffmann, K., Drabkina, M., Arampatzi, P., et al. (2021). Opposing Wnt signals regulate cervical squamocolumnar homeostasis and emergence of metaplasia. Nat. Cell Biol. *23*, 184–197. https://doi.org/10.1038/s41556-020-00619-0.

50. Cooley, A., Madhukaran, S., Stroebele, E., Colon Caraballo, M., Wang, L., Akgul, Y., Hon, G.C., and Mahendroo, M. (2023). Dynamic states of cervical epithelia during pregnancy and epithelial barrier disruption. iScience *26*, 105953. https://doi.org/10.1016/j.isci.2023.105953.

51. Hauser, B.R., Aure, M.H., Kelly, M.C., Genomics and Computational Biology Core, Hoffman, M.P., and Chibly, A.M. (2020). Generation of a Single-Cell RNAseq Atlas of Murine Salivary Gland Development. iScience *23*, 101838. https://doi.org/10.1016/j.isci.2020.101838.

52. Horeth, E., Oyelakin, A., Song, E.-A.C., Che, M., Bard, J., Min, S., Kiripolsky, J., Kramer, J.M., Sinha, S., and Romano, R.-A. (2021). Transcriptomic and Single-Cell Analysis Reveals Regulatory Networks and Cellular Heterogeneity in Mouse Primary Sjögren's Syndrome Salivary Glands. Front. Immunol. *12*, 729040. https://doi.org/10.3389/fimmu.2021.729040.

53. Yu, Z., Liao, J., Chen, Y., Zou, C., Zhang, H., Cheng, J., Liu, D., Li, T., Zhang, Q., Li, J., et al. (2019). Single-Cell Transcriptomic Map of the Human and Mouse Bladders. J. Am. Soc. Nephrol. JASN *30*, 2159–2176. https://doi.org/10.1681/ASN.2019040335.

54. Baker, D., Al-Naggar, I.M., Sivajothi, S., Flynn, W.F., Amiri, A., Luo, D., Hardy, C.C., Kuchel, G.A., Smith, P.P., and Robson, P. (2021). A Cellular Reference Resource for the Mouse Urinary Bladder. Preprint at bioRxiv, https://doi.org/10.1101/2021.09.20.461121 https://doi.org/10.1101/2021.09.20.461121.

55. Kim, E., Choi, S., Kang, B., Kong, J., Kim, Y., Yoon, W.H., Lee, H.-R., Kim, S., Kim, H.-M., Lee, H., et al. (2020). Creation of bladder assembloids mimicking tissue regeneration and cancer. Nature *588*, 664–669. https://doi.org/10.1038/s41586-020-3034-x.

56. Li, Y., Liu, Y., Gao, Z., Zhang, L., Chen, L., Wu, Z., Liu, Q., Wang, S., Zhou, N., Chai, T.C., et al. (2021). Single-cell transcriptomes of mouse bladder urothelium uncover novel cell type markers and urothelial differentiation characteristics. Cell Prolif. *54*, e13007. https://doi.org/10.1111/cpr.13007.

57. Montoro, D.T., Haber, A.L., Biton, M., Vinarsky, V., Lin, B., Birket, S.E., Yuan, F., Chen, S., Leung, H.M., Villoria, J., et al. (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature *560*, 319–324. https://doi.org/10.1038/s41586-018-0393-7.

58. Kiyokawa, H., Yamaoka, A., Matsuoka, C., Tokuhara, T., Abe, T., and Morimoto, M. (2021). Airway basal stem cells reutilize the embryonic proliferation regulator, Tgfβ-Id2 axis, for tissue regeneration. Dev. Cell *56*, 1917-1929.e9. https://doi.org/10.1016/j.devcel.2021.05.016.

59. Lin, B., Shah, V.S., Chernoff, C., Sun, J., Shipkovenska, G.G., Vinarsky, V., Waghray, A., Xu, J., Leduc, A.D., Hintschich, C.A., et al. (2024). Airway hillocks are injury-resistant reservoirs of unique plastic stem cells. Nature *629*, 869–877. https://doi.org/10.1038/s41586-024-07377-1.

60. Huang, D., Jiao, X., Huang, S., Liu, J., Si, H., Qi, D., Pei, X., Lu, D., Wang, Y., and Li, Z. (2024). Analysis of the heterogeneity and complexity of murine extraorbital lacrimal gland via single-cell RNA sequencing. Ocul. Surf. *34*, 60–95. https://doi.org/10.1016/j.jtos.2024.06.005.

61. Delcroix, V., Mauduit, O., Lee, H.S., Ivanova, A., Umazume, T., Knox, S.M., de Paiva, C.S., Dartt, D.A., and Makarenkova, H.P. (2023). The First Transcriptomic Atlas of the Adult Lacrimal Gland Reveals Epithelial Complexity and Identifies Novel Progenitor Cells in Mice. Cells *12*, 1435. https://doi.org/10.3390/cells12101435.

62. Rattner, A., Heng, J.S., Winer, B.L., Goff, L.A., and Nathans, J. (2023). Normal and Sjogren's syndrome models of the murine lacrimal gland studied at single-cell resolution. Proc. Natl. Acad. Sci. U. S. A. *120*, e2311983120. https://doi.org/10.1073/pnas.2311983120.

63. Vercauteren Drubbel, A., and Beck, B. (2023). Single-cell transcriptomics uncovers the differentiation of a subset of murine esophageal progenitors into taste buds in vivo. Sci. Adv. *9*, eadd9135. https://doi.org/10.1126/sciadv.add9135.

64. Kabir, M.F., Karami, A.L., Cruz-Acuña, R., Klochkova, A., Saxena, R., Mu, A., Murray, M.G., Cruz, J., Fuller, A.D., Clevenger, M.H., et al. (2022). Single cell transcriptomic analysis reveals cellular diversity of murine esophageal epithelium. Nat. Commun. *13*, 2167. https://doi.org/10.1038/s41467-022-29747-x.

65. McGinn, J., Hallou, A., Han, S., Krizic, K., Ulyanchenko, S., Iglesias-Bartolome, R., England, F.J., Verstreken, C., Chalut, K.J., Jensen, K.B., et al. (2021). A biomechanical switch regulates the transition towards homeostasis in oesophageal epithelium. Nat. Cell Biol. *23*, 511–525. https://doi.org/10.1038/s41556-021-00679-w.

66. Karthaus, W.R., Hofree, M., Choi, D., Linton, E.L., Turkekul, M., Bejnood, A., Carver, B., Gopalan, A., Abida, W., Laudone, V., et al. (2020). Regenerative potential of prostate luminal cells revealed by single-cell analysis. Science *368*, 497–505. https://doi.org/10.1126/science.aay0267.

67. Crowley, L., Cambuli, F., Aparicio, L., Shibata, M., Robinson, B.D., Xuan, S., Li, W., Hibshoosh, H., Loda, M., Rabadan, R., et al. (2020). A single-cell atlas of the mouse and human prostate reveals heterogeneity and conservation of epithelial progenitors. eLife *9*, e59465. https://doi.org/10.7554/eLife.59465.

68. Aparicio, L., Crowley, L., Christin, J.R., Laplaca, C.J., Hibshoosh, H., Rabadan, R., and Shen, M.M. (2025). Meta-analyses of mouse and human prostate single-cell transcriptomes reveal widespread epithelial plasticity in tissue regression, regeneration, and cancer. Genome Med. *17*, 5. https://doi.org/10.1186/s13073-025-01432-w.

69. Graham, M.K., Chikarmane, R., Wang, R., Vaghasia, A., Gupta, A., Zheng, Q., Wodu, B., Pan, X., Castagna, N., Liu, J., et al. (2023). Single-cell atlas of epithelial and stromal cell heterogeneity by lobe and strain in the mouse prostate. The Prostate *83*, 286–303. https://doi.org/10.1002/pros.24460.

70. Joseph, D.B., Henry, G.H., Malewska, A., Iqbal, N.S., Ruetten, H.M., Turco, A.E., Abler, L.L., Sandhu, S.K., Cadena, M.T., Malladi, V.S., et al. (2020). Urethral luminal epithelia are castration-insensitive cells of the proximal prostate. The Prostate *80*, 872–884. https://doi.org/10.1002/pros.24020.

71. Joseph, D.B., Henry, G.H., Malewska, A., Reese, J.C., Mauck, R.J., Gahan, J.C., Hutchinson, R.C., Malladi, V.S., Roehrborn, C.G., Vezina, C.M., et al. (2021). Single-cell analysis of mouse and human prostate reveals novel fibroblasts with specialized distribution and microenvironment interactions. J. Pathol. *255*, 141–154. https://doi.org/10.1002/path.5751.

72. He, Z., Brazovskaja, A., Ebert, S., Camp, J.G., and Treutlein, B. (2020). CSS: cluster similarity spectrum integration of single-cell genomics data. Genome Biol. *21*, 224. https://doi.org/10.1186/s13059-020-02147-4.

73. Mistry, D.S., Chen, Y., Wang, Y., and Sen, G.L. (2014). SNAI2 Controls the Undifferentiated State of Human Epidermal Progenitor Cells. Stem Cells Dayt. Ohio *32*, 3209–3218. https://doi.org/10.1002/stem.1809.

74. Fan, X., Wang, D., Burgmaier, J.E., Teng, Y., Romano, R.-A., Sinha, S., and Yi, R. (2018). Single Cell and Open Chromatin Analysis Reveals Molecular Origin of Epidermal Cells of the Skin. Dev. Cell *47*, 21-37.e5. https://doi.org/10.1016/j.devcel.2018.08.010.

75. Richardson, R.J., Dixon, J., Malhotra, S., Hardman, M.J., Knowles, L., Boot-Handford, R.P., Shore, P., Whitmarsh, A., and Dixon, M.J. (2006). Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch. Nat. Genet. *38*, 1329–1334. https://doi.org/10.1038/ng1894.

76. Gordon, W.M., Zeller, M.D., Klein, R.H., Swindell, W.R., Ho, H., Espetia, F., Gudjonsson, J.E., Baldi, P.F., and Andersen, B. (2014). A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. J. Clin. Invest. *124*, 5205–5218. https://doi.org/10.1172/JCI77138.

77. Burkly, L., Hession, C., Ogata, L., Reilly, C., Marconi, L.A., Olson, D., Tizard, R., Cate, R., and Lo, D. (1995). Expression of relB is required for the development of thymic medulla and dendritic cells. Nature *373*, 531–536. https://doi.org/10.1038/373531a0.

78. Michelson, D.A., Hase, K., Kaisho, T., Benoist, C., and Mathis, D. (2022). Thymic epithelial cells co-opt lineage-defining transcription factors to eliminate autoreactive T cells. Cell *185*, 2542-2558.e18. https://doi.org/10.1016/j.cell.2022.05.018.

79. Dempsey, L.A. (2023). NF-κB2 defects in autoimmunity. Nat. Immunol. *24*, 1965–1965. https://doi.org/10.1038/s41590-023-01701-z.

80. Qin, J., Wu, S.-P., Creighton, C.J., Dai, F., Xie, X., Cheng, C.-M., Frolov, A., Ayala, G., Lin, X., Feng, X.-H., et al. (2013). COUP-TFII inhibits TGF-β-induced growth barrier to promote prostate tumorigenesis. Nature *493*, 236–240. https://doi.org/10.1038/nature11674.

81. Chan, C.-P., Kok, K.-H., and Jin, D.-Y. (2011). CREB3 subfamily transcription factors are not created equal: Recent insights from global analyses and animal models. Cell Biosci. *1*, 6. https://doi.org/10.1186/2045-3701-1-6.

82. Ehebauer, M., Hayward, P., and Arias, A.M. (2006). Notch, a Universal Arbiter of Cell Fate Decisions. Science *314*, 1414–1415. https://doi.org/10.1126/science.1134042.

83. Fre, S., Hannezo, E., Sale, S., Huyghe, M., Lafkas, D., Kissel, H., Louvi, A., Greve, J., Louvard, D., and Artavanis-Tsakonas, S. (2011). Notch Lineages and Activity in Intestinal Stem Cells Determined by a New Set of Knock-In Mice. PLoS ONE *6*, e25785. https://doi.org/10.1371/journal.pone.0025785.

84. Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L., and Luo, L. (2007). A global double-fluorescent Cre reporter mouse. genesis *45*, 593–605. https://doi.org/10.1002/dvg.20335.

85. Barbieri, C.E., Perez, C.A., Johnson, K.N., Ely, K.A., Billheimer, D., and Pietenpol, J.A. (2005). IGFBP-3 Is a Direct Target of Transcriptional Regulation by ΔNp63α in Squamous Epithelium. Cancer Res. *65*, 2314–2320. https://doi.org/10.1158/0008-5472.CAN-04-3449.

86. Ihrie, R.A., Marques, M.R., Nguyen, B.T., Horner, J.S., Papazoglu, C., Bronson, R.T., Mills, A.A., and Attardi, L.D. (2005). Perp Is a p63-Regulated Gene Essential for Epithelial Integrity. Cell *120*, 843–856. https://doi.org/10.1016/j.cell.2005.01.008.

87. Kawata, M., Taniguchi, Y., Mori, D., Yano, F., Ohba, S., Chung, U., Shimogori, T., Mills, A.A., Tanaka, S., and Saito, T. (2017). Different regulation of limb development by p63 transcript variants. PLOS ONE *12*, e0174122. https://doi.org/10.1371/journal.pone.0174122.

88. Belyi, V.A., Ak, P., Markert, E., Wang, H., Hu, W., Puzio-Kuter, A., and Levine, A.J. (2010). The Origins and Evolution of the p53 Family of Genes. Cold Spring Harb. Perspect. Biol. *2*, a001198. https://doi.org/10.1101/cshperspect.a001198.

89. Li, H., Janssens, J., De Waegeneer, M., Kolluru, S.S., Davie, K., Gardeux, V., Saelens, W., David, F.P.A., Brbić, M., Spanier, K., et al. (2022). Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. Science *375*, eabk2432. https://doi.org/10.1126/science.abk2432.

90. Lachner, J., Ehrlich, F., Wielscher, M., Farlik, M., Hermann, M., Tschachler, E., and Eckhart, L. (2022). Single-cell transcriptomics defines keratinocyte differentiation in avian scutate scales. Sci. Rep. *12*, 126. https://doi.org/10.1038/s41598-021-04082-1.

91. Gao, Y., Yao, X., Zhai, Y., Li, L., Li, H., Sun, X., Yu, P., Xue, T., Li, Y., and Hu, Y. (2021). Single cell transcriptional zonation of human psoriasis skin identifies an alternative immunoregulatory axis conducted by skin resident cells. Cell Death Dis. *12*, 450. https://doi.org/10.1038/s41419-021-03724-6.

92. Pang, Y., Qin, Y., Du, Z., Liu, Q., Zhang, J., Han, K., Lu, J., Yuan, Z., Li, J., Pan, S., et al. (2025). Single-cell transcriptome atlas of lamprey exploring Natterin- induced white adipose tissue browning. Nat. Commun. *16*, 752. https://doi.org/10.1038/s41467-025-56153-w.

93. Liao, Y., Ma, L., Guo, Q., E, W., Fang, X., Yang, L., Ruan, F., Wang, J., Zhang, P., Sun, Z., et al. (2022). Cell landscape of larval and adult Xenopus laevis at single-cell resolution. Nat. Commun. *13*, 4306. https://doi.org/10.1038/s41467-022-31949-2.

94. Aman, A.J., Saunders, L.M., Carr, A.A., Srivatasan, S., Eberhard, C., Carrington, B., Watkins-Chow, D., Pavan, W.J., Trapnell, C., and Parichy, D.M. Transcriptomic profiling of tissue environments critical for post-embryonic patterning and morphogenesis of zebrafish skin. eLife *12*, RP86670. https://doi.org/10.7554/eLife.86670.

95. Akat, E., Yenmiş, M., Pombal, M.A., Molist, P., Megías, M., Arman, S., Veselỳ, M., Anderson, R., and Ayaz, D. (2022). Comparison of vertebrate skin structure at class level: A review. Anat. Rec. *305*, 3543–3608. https://doi.org/10.1002/ar.24908.

96. Drost, H.-G., Gabel, A., Liu, J., Quint, M., and Grosse, I. (2018). myTAI: evolutionary transcriptomics with R. Bioinformatics *34*, 1589–1590. https://doi.org/10.1093/bioinformatics/btx835.

97. Fessing, M.Y., Mardaryev, A.N., Gdula, M.R., Sharov, A.A., Sharova, T.Y., Rapisarda, V., Gordon, K.B., Smorodchenko, A.D., Poterlowicz, K., Ferone, G., et al. (2011). p63

regulates Satb1 to control tissue-specific chromatin remodeling during development of the epidermis. J. Cell Biol. *194*, 825–839. https://doi.org/10.1083/jcb.201101148.

98. Mardaryev, A.N., Gdula, M.R., Yarker, J.L., Emelianov, V.U., Poterlowicz, K., Sharov, A.A., Sharova, T.Y., Scarpa, J.A., Joffe, B., Solovei, I., et al. (2014). p63 and Brg1 control developmentally regulated higher-order chromatin remodelling at the epidermal differentiation complex locus in epidermal progenitor cells. Dev. Camb. Engl. *141*, 101–111. https://doi.org/10.1242/dev.103200.

99. Yu, Z., Lin, K.K., Bhandari, A., Spencer, J.A., Xu, X., Wang, N., Lu, Z., Gill, G.N., Roop, D.R., Wertz, P., et al. (2006). The Grainyhead-like epithelial transactivator Get-1/Grhl3 regulates epidermal terminal differentiation and interacts functionally with LMO4. Dev. Biol. *299*, 122–136. https://doi.org/10.1016/j.ydbio.2006.07.015.

100. Chen, W., Xiao Liu, Z., Oh, J.-E., Shin, K.-H., Kim, R.H., Jiang, M., Park, N.-H., and Kang, M.K. (2012). Grainyhead-like 2 (GRHL2) inhibits keratinocyte differentiation through epigenetic mechanism. Cell Death Dis. *3*, e450–e450. https://doi.org/10.1038/cddis.2012.190.

101. Sethi, I., Gluck, C., Zhou, H., Buck, M.J., and Sinha, S. (2017). Evolutionary re-wiring of p63 and the epigenomic regulatory landscape in keratinocytes and its potential implications on species-specific gene expression and phenotypes. Nucleic Acids Res. *45*, 8208–8224. https://doi.org/10.1093/nar/gkx416.

102. Klein, R.H., Lin, Z., Hopkin, A.S., Gordon, W., Tsoi, L.C., Liang, Y., Gudjonsson, J.E., and Andersen, B. (2017). GRHL3 binding and enhancers rearrange as epidermal keratinocytes transition between functional states. PLOS Genet. *13*, e1006745. https://doi.org/10.1371/journal.pgen.1006745.

103. Ting, S.B., Caddy, J., Hislop, N., Wilanowski, T., Auden, A., Zhao, L., Ellis, S., Kaur, P., Uchida, Y., Holleran, W.M., et al. (2005). A Homolog of Drosophila grainy head Is Essential for Epidermal Integrity in Mice. Science *308*, 411–413. https://doi.org/10.1126/science.1107511.

104. Hu, C., Diévart, A., Lupien, M., Calvo, E., Tremblay, G., and Jolicoeur, P. (2006). Overexpression of Activated Murine Notch1 and Notch3 in Transgenic Mice Blocks Mammary Gland Development and Induces Mammary Tumors. Am. J. Pathol. *168*, 973–990. https://doi.org/10.2353/ajpath.2006.050416.

105. Gouilleux, F., Pallard, C., Dusanter-Fourt, I., Wakao, H., Haldosen, L.A., Norstedt, G., Levy, D., and Groner, B. (1995). Prolactin, growth hormone, erythropoietin and granulocyte-macrophage colony stimulating factor induce MGF-Stat5 DNA binding activity. EMBO J. *14*, 2005–2013. https://doi.org/10.1002/j.1460-2075.1995.tb07192.x.

106. Koster, M.I. (2010). p63 in skin development and ectodermal dysplasias. J. Invest. Dermatol. *130*, 2352–2358. https://doi.org/10.1038/jid.2010.119.

107. McKeon, F. (2004). p63 and the epithelial stem cell: more than status quo? Genes Dev. *18*, 465–469. https://doi.org/10.1101/gad.1190504.

108. Tadeu, A.M.B., and Horsley, V. (2013). Notch signaling represses p63 expression in the developing surface ectoderm. Dev. Camb. Engl. *140*, 3777–3786. https://doi.org/10.1242/dev.093948.

109. De Rosa, L., Antonini, D., Ferone, G., Russo, M.T., Yu, P.B., Han, R., and Missero, C. (2009). p63 Suppresses non-epidermal lineage markers in a bone morphogenetic protein-dependent manner via repression of Smad7. J. Biol. Chem. *284*, 30574–30582. https://doi.org/10.1074/jbc.M109.049619.

110. Santos-Pereira, J.M., Gallardo-Fuentes, L., Neto, A., Acemel, R.D., and Tena, J.J. (2019). Pioneer and repressive functions of p63 during zebrafish embryonic ectoderm specification. Nat. Commun. *10*, 3049. https://doi.org/10.1038/s41467-019-11121-z.

111. Shalom-Feuerstein, R., Lena, A.M., Zhou, H., De La Forest Divonne, S., Van Bokhoven, H., Candi, E., Melino, G., and Aberdam, D. (2011). ΔNp63 is an ectodermal gatekeeper of epidermal morphogenesis. Cell Death Differ. *18*, 887–896. https://doi.org/10.1038/cdd.2010.159.

112. Medawar, A., Virolle, T., Rostagno, P., Forest-Divonne, S. de la, Gambaro, K., Rouleau, M., and Aberdam, D. (2008). ΔNp63 Is Essential for Epidermal Commitment of Embryonic Stem Cells. PLOS ONE *3*, e3441. https://doi.org/10.1371/journal.pone.0003441.

113. Aberdam, E., Barak, E., Rouleau, M., de LaForest, S., Berrih-Aknin, S., Suter, D.M., Krause, K.-H., Amit, M., Itskovitz-Eldor, J., and Aberdam, D. (2008). A pure population of ectodermal cells derived from human embryonic stem cells. Stem Cells Dayt. Ohio *26*, 440–444. https://doi.org/10.1634/stemcells.2007-0588.

114. Barton, C.E., Tahinci, E., Barbieri, C.E., Johnson, K.N., Hanson, A.J., Jernigan, K.K., Chen, T.W., Lee, E., and Pietenpol, J.A. (2009). ΔNp63 antagonizes p53 to regulate mesoderm induction in Xenopus laevis. Dev. Biol. *329*, 130–139. https://doi.org/10.1016/j.ydbio.2009.02.036.

115. Carroll, S.B. (2008). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell *134*, 25–36. https://doi.org/10.1016/j.cell.2008.06.030.

116. Jacob, F. (1977). Evolution and Tinkering. Science *196*, 1161–1166. https://doi.org/10.1126/science.860134.

117. Blanpain, C., and Fuchs, E. (2009). Epidermal homeostasis: a balancing act of stem cells in the skin. Nat. Rev. Mol. Cell Biol. *10*, 207–217. https://doi.org/10.1038/nrm2636.

118. Ali, A., Syed, S.M., Jamaluddin, M.F.B., Colino-Sanguino, Y., Gallego-Ortega, D., and Tanwar, P.S. (2020). Cell Lineage Tracing Identifies Hormone-Regulated and Wnt-Responsive Vaginal Epithelial Stem Cells. Cell Rep. *30*, 1463-1477.e7. https://doi.org/10.1016/j.celrep.2020.01.003.

119. Doupé, D.P., Alcolea, M.P., Roshan, A., Zhang, G., Klein, A.M., Simons, B.D., and Jones, P.H. (2012). A Single Progenitor Population Switches Behavior to Maintain and Repair Esophageal Epithelium. Science *337*, 1091–1093. https://doi.org/10.1126/science.1218835.

120. Chatzeli, L., Bordeu, I., Han, S., Bisetto, S., Waheed, Z., Koo, B.-K., Alcolea, M.P., and Simons, B.D. (2023). A cellular hierarchy of Notch and Kras signaling controls cell fate specification in the developing mouse salivary gland. Dev. Cell *58*, 94-109.e6. https://doi.org/10.1016/j.devcel.2022.12.009.

121. Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*, 189–193. https://doi.org/10.1038/nature10573.

122. Farmer, D.T., Nathan, S., Finley, J.K., Shengyang Yu, K., Emmerson, E., Byrnes, L.E., Sneddon, J.B., McManus, M.T., Tward, A.D., and Knox, S.M. (2017). Defining epithelial cell dynamics and lineage relationships in the developing lacrimal gland. Dev. Camb. Engl. *144*, 2517–2528. https://doi.org/10.1242/dev.150789.

123. Choi, N., Zhang, B., Zhang, L., Ittmann, M., and Xin, L. (2012). Adult murine prostate basal and luminal cells are self-sustained lineages that can both serve as targets for prostate cancer initiation. Cancer Cell *21*, 253–265. https://doi.org/10.1016/j.ccr.2012.01.005.

124. Ousset, M., Van Keymeulen, A., Bouvencourt, G., Sharma, N., Achouri, Y., Simons, B.D., and Blanpain, C. (2012). Multipotent and unipotent progenitors contribute to prostate postnatal development. Nat. Cell Biol. *14*, 1131–1138. https://doi.org/10.1038/ncb2600.

125. Centonze, A., Lin, S., Tika, E., Sifrim, A., Fioramonti, M., Malfait, M., Song, Y., Wuidart, A., Van Herck, J., Dannau, A., et al. (2020). Heterotypic cell-cell communication regulates glandular stem cell multipotency. Nature *584*, 608–613. https://doi.org/10.1038/s41586-020-2632-y.

126. Nam, Y., Sliz, P., Pear, W.S., Aster, J.C., and Blacklow, S.C. (2007). Cooperative assembly of higher-order Notch complexes functions as a switch to induce transcription. Proc. Natl. Acad. Sci. *104*, 2103–2108. https://doi.org/10.1073/pnas.0611092104.

127. Bray, S.J. (2006). Notch signalling: a simple pathway becomes complex. Nat. Rev. Mol. Cell Biol. *7*, 678–689. https://doi.org/10.1038/nrm2009.

128. Parker, J., and Pennell, M. (2025). The cellular substrate of evolutionary novelty. Curr. Biol. CB *35*, R626–R637. https://doi.org/10.1016/j.cub.2025.04.014.

129. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

130. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

131. Xu, S., Hu, E., Cai, Y., Xie, Z., Luo, X., Zhan, L., Tang, W., Wang, Q., Liu, B., Wang, R., et al. (2024). Using clusterProfiler to characterize multiomics data. Nat. Protoc. *19*, 3292–3320. https://doi.org/10.1038/s41596-024-01020-z.

132. Su, C., Xu, Z., Shan, X., Cai, B., Zhao, H., and Zhang, J. (2023). Cell-type-specific co-expression inference from single cell RNA-sequencing data. Nat. Commun. *14*, 4846. https://doi.org/10.1038/s41467-023-40503-7.

133. Barrera-Redondo, J., Lotharukpong, J.S., Drost, H.-G., and Coelho, S.M. (2023). Uncovering gene-family founder events during major evolutionary transitions in animals, plants and fungi using GenEra. Genome Biol. *24*, 54. https://doi.org/10.1186/s13059-023-02895-z.

134. Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., and Kasprzyk, A. (2009). BioMart – biological queries made easy. BMC Genomics *10*, 22. https://doi.org/10.1186/1471-2164-10-22.

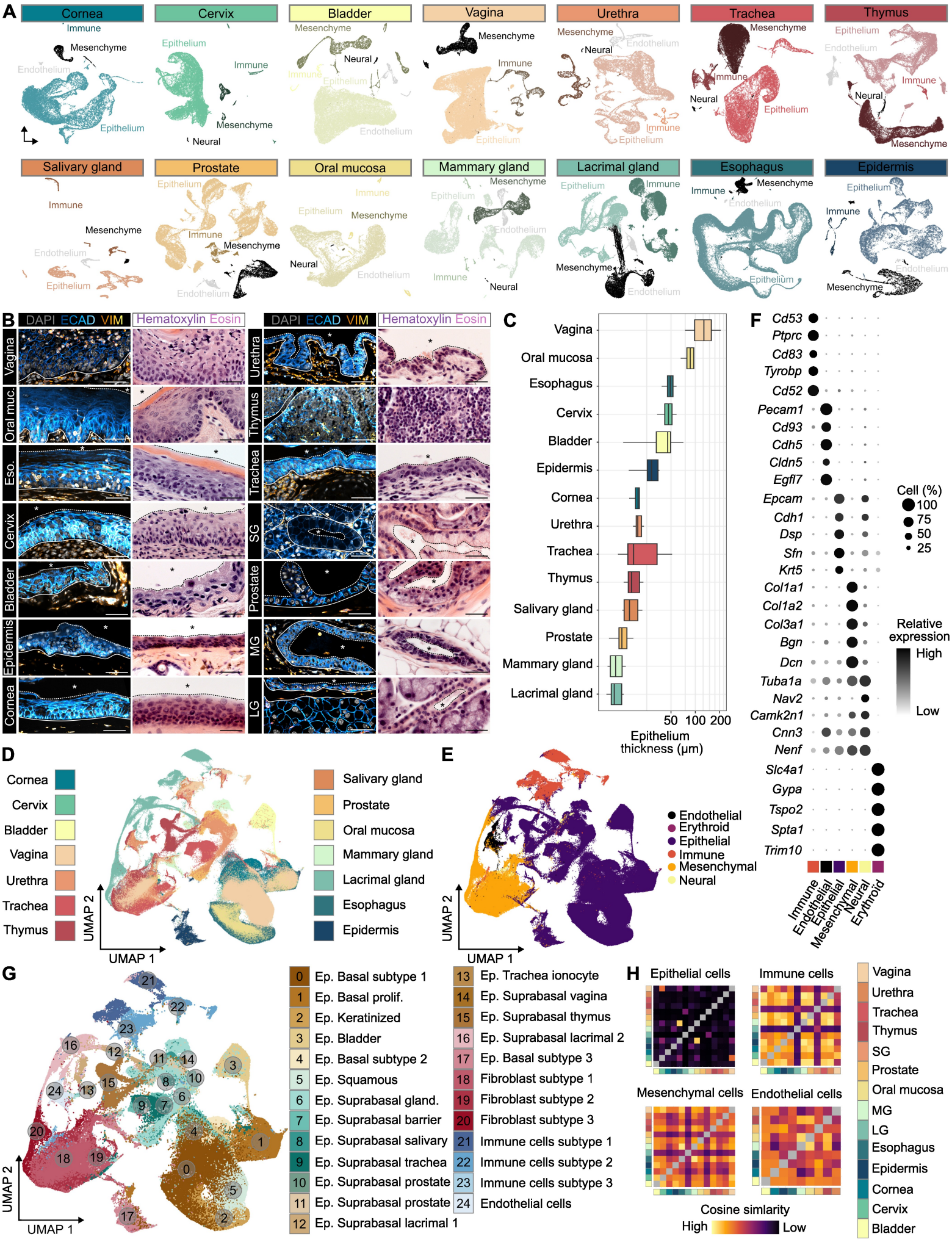

A
Cornea
Immune
Mesenchyme
Endothelium
Epithelium
Cervix
Epithelium
Immune
Mesenchyme
Neural
Bladder
Mesenchyme
Neural
Immune
Epithelium
Endothelium
Vagina
Mesenchyme
Immune
Epithelium
Urethra
Mesenchyme
Endothelium
Neural
Immune
Epithelium
Trachea
Mesenchyme
Immune
Neural
Epithelium
Thymus
Epithelium
Endothelium
Immune
Neural
Mesenchyme
Salivary gland
Immune
Mesenchyme
Endothelium
Epithelium
Prostate
Epithelium
Immune
Mesenchyme
Endothelium
Oral mucosa
Immune
Epithelium
Mesenchyme
Neural
Endothelium
Mammary gland
Mesenchyme
Epithelium
Endothelium
Immune
Neural
Lacrimal gland
Epithelium
Immune
Mesenchyme
Endothelium
Esophagus
Immune
Mesenchyme
Endothelium
Epithelium
Epidermis
Epithelium
Immune
Endothelium
Mesenchyme
B
DAPI ECAD VIM
Hematoxylin Eosin
Vagina
Oral muc.
Eso.
Cervix
Bladder
Epidermis
Cornea
Urethra
Thymus
Trachea
SG
Prostate
MG
LG
C
Vagina
Oral mucosa
Esophagus
Cervix
Bladder
Epidermis
Cornea
Urethra
Trachea
Thymus
Salivary gland
Prostate
Mammary gland
Lacrimal gland
50 100 200
Epithelium thickness (µm)
F
Cd53
Ptprc
Cd83
Tyrobp
Cd52
Pecam1
Cd93
Cdh5
Cldn5
Egfl7
Epcam
Cdh1
Dsp
Sfn
Krt5
Col1a1
Col1a2
Col3a1
Bgn
Dcn
Tuba1a
Nav2
Camk2n1
Cnn3
Nenf
Slc4a1
Gypa
Tspo2
Spta1
Trim10
Cell (%)
100
75
50
25
Relative expression
High
Low
Immune
Endothelial
Epithelial
Mesenchymal
Neural
Erythroid
D
Cornea
Cervix
Bladder
Vagina
Urethra
Trachea
Thymus
Salivary gland
Prostate
Oral mucosa
Mammary gland
Lacrimal gland
Esophagus
Epidermis
UMAP 2
UMAP 1
E
Endothelial
Erythroid
Epithelial
Immune
Mesenchymal
Neural
UMAP 2
UMAP 1
G
UMAP 2
UMAP 1
0 Ep. Basal subtype 1
1 Ep. Basal prolif.
2 Ep. Keratinized
3 Ep. Bladder
4 Ep. Basal subtype 2
5 Ep. Squamous
6 Ep. Suprabasal gland.
7 Ep. Suprabasal barrier
8 Ep. Suprabasal salivary
9 Ep. Suprabasal trachea
10 Ep. Suprabasal prostate
11 Ep. Suprabasal prostate
12 Ep. Suprabasal lacrimal 1
13 Ep. Trachea ionocyte
14 Ep. Suprabasal vagina
15 Ep. Suprabasal thymus
16 Ep. Suprabasal lacrimal 2
17 Ep. Basal subtype 3
18 Fibroblast subtype 1
19 Fibroblast subtype 2
20 Fibroblast subtype 3
21 Immune cells subtype 1
22 Immune cells subtype 2
23 Immune cells subtype 3
24 Endothelial cells
H
Epithelial cells
Immune cells
Mesenchymal cells
Endothelial cells
Cosine similarity
High
Low
Vagina
Urethra
Trachea
Thymus
SG
Prostate
Oral mucosa
MG
LG
Esophagus
Epidermis
Cornea
Cervix
Bladder

Figure 1. Integrated transcriptomic cell atlas of murine multilayered epithelia organs

A: *Uniform manifold approximation and projection (UMAP) of scRNA-seq data colored by organ and cell class.* ***B:*** *Confocal imaging of adult WT multilayered epithelia stained for Dapi, Ecadherin (ECAD, Cyan) and Vimentin (VIM, Orange) or Hematoxylin & Eosin staining. *: indicate lumen or external space; dotted line: indicate the interface between apical domain of epithelia and external space; continuous line: indicate the interface between epithelium and basement membrane. Scale bar: 50 µm.* ***C:*** *Quantification of epithelium thickness across multilayered epithelia (n=10 independent measurements).* ***D-E:*** *Integrated UMAP of scRNA-seq data colored by organ (D), and cell class (E).* ***F:*** *Dot plot showing cell class-level marker gene expression.* ***G:*** *UMAP of the multi-organ atlas, colored by major cell cluster, with shade and numbers representing cluster assignments.* ***H:*** *Cosine similarity between transcriptomes across organs, stratified by major cell class. To ensure robust representation, a threshold (≥100 cells) was applied for Immune, Mesenchymal, and Endothelial cell classes, retaining only tissues meeting this criterion. Rows and columns are color coded based on tissues.*

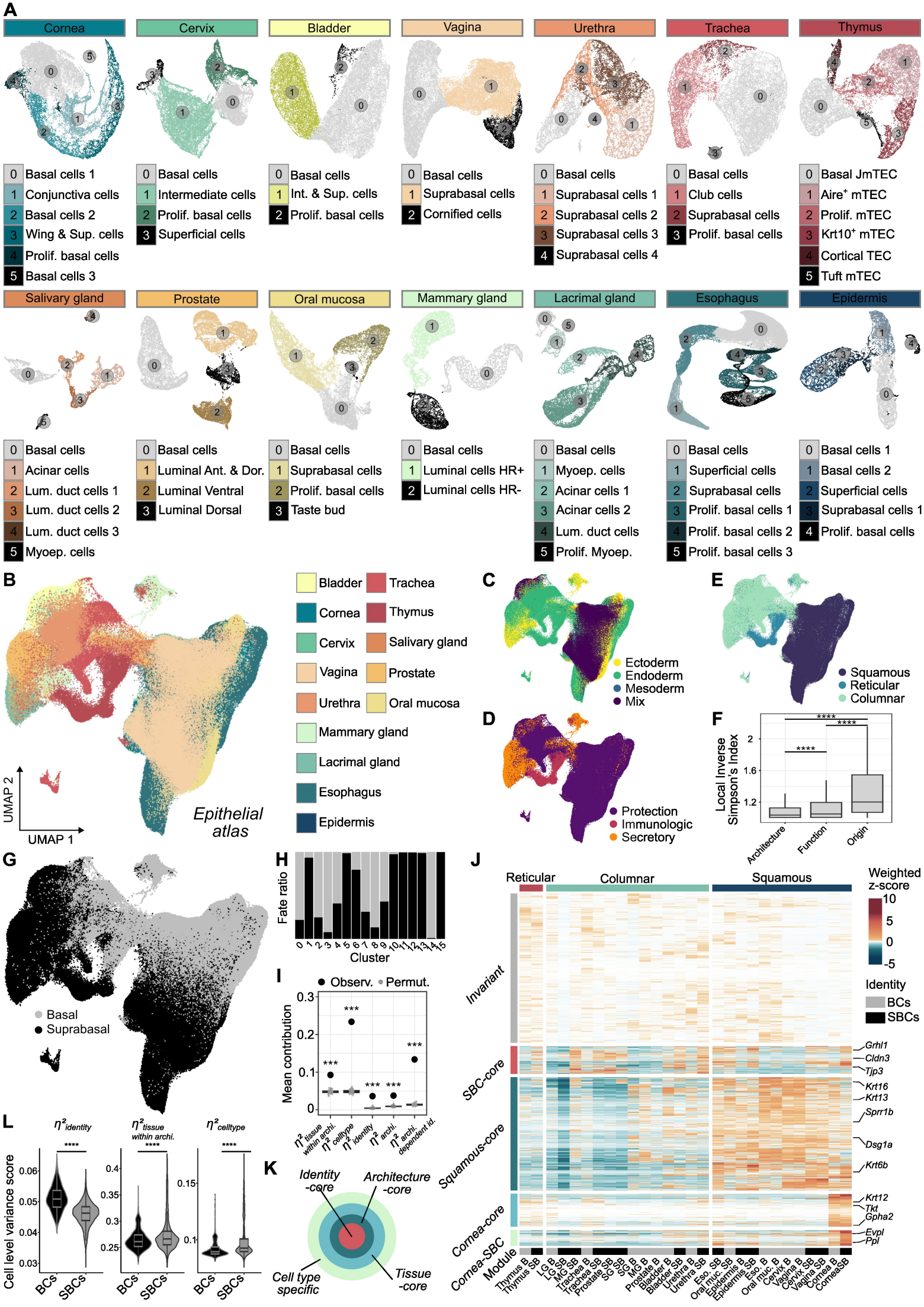

A
Cornea
0 Basal cells 1
1 Conjunctiva cells
2 Basal cells 2
3 Wing & Sup. cells
4 Prolif. basal cells
5 Basal cells 3
Cervix
0 Basal cells
1 Intermediate cells
2 Prolif. basal cells
3 Superficial cells
Bladder
0 Basal cells
1 Int. & Sup. cells
2 Prolif. basal cells
Vagina
0 Basal cells
1 Suprabasal cells
2 Cornified cells
Urethra
0 Basal cells
1 Suprabasal cells 1
2 Suprabasal cells 2
3 Suprabasal cells 3
4 Suprabasal cells 4
Trachea
0 Basal cells
1 Club cells
2 Suprabasal cells
3 Prolif. basal cells
Thymus
0 Basal JmTEC
1 Aire+ mTEC
2 Prolif. mTEC
3 Krt10+ mTEC
4 Cortical TEC
5 Tuft mTEC
Salivary gland
0 Basal cells
1 Acinar cells
2 Lum. duct cells 1
3 Lum. duct cells 2
4 Lum. duct cells 3
5 Myoep. cells
Prostate
0 Basal cells
1 Luminal Ant. & Dor.
2 Luminal Ventral
3 Luminal Dorsal
Oral mucosa
0 Basal cells
1 Suprabasal cells
2 Prolif. basal cells
3 Taste bud
Mammary gland
0 Basal cells
1 Luminal cells HR+
2 Luminal cells HR-
Lacrimal gland
0 Basal cells
1 Myoep. cells
2 Acinar cells 1
3 Acinar cells 2
4 Lum. duct cells
5 Prolif. Myoep.
Esophagus
0 Basal cells
1 Superficial cells
2 Suprabasal cells
3 Prolif. basal cells 1
4 Prolif. basal cells 2
5 Prolif. basal cells 3
Epidermis
0 Basal cells 1
1 Basal cells 2
2 Superficial cells
3 Suprabasal cells 1
4 Prolif. basal cells
B
UMAP 2
UMAP 1
Epithelial atlas
Bladder
Cornea
Cervix
Vagina
Urethra
Mammary gland
Lacrimal gland
Esophagus
Epidermis
Trachea
Thymus
Salivary gland
Prostate
Oral mucosa
C
Ectoderm
Endoderm
Mesoderm
Mix
D
Protection
Immunologic
Secretory
E
Squamous
Reticular
Columnar
F
Local Inverse Simpson's Index
Architecture
Function
Origin
G
Basal
Suprabasal
H
Fate ratio
Cluster
I
Observ.
Permut.
Mean contribution
J
Reticular
Columnar
Squamous
Weighted z-score
Identity
BCs
SBCs
Invariant
SBC-core
Squamous-core
Cornea-core
Cornea-SBC
Module
Grhl1
Cldn3
Tjp3
Krt16
Krt13
Sprr1b
Dsg1a
Krt6b
Krt12
Tkt
Gpha2
Evpl
Ppl
K
Identity -core
Architecture -core
Cell type specific
Tissue -core
L
Cell level variance score
BCs
SBCs

Figure 2. Identification of conserved spatial and transcriptomic organization in multilayered epithelia.

A: *UMAP of epithelial cells scRNA-seq data colored by organ and cell cluster. Prolif.: Proliferative cells; Sup.: Suprabasal cells; mTEC: medullar Thymic epithelial cells; Lum.: Luminal; Myoep.: Myoepithelial; Ant.: Anterior; Dor.: Dorsal; HR: Hormone responsive.* ***B-E:*** *Integrated UMAP of epithelial scRNAseq data colored by organ (B), embryonic origin (C), physiological function (D) and tissue architecture (E).* ***F:*** *Local Inverse Simpson's Index (LISI) scores for the epithelial integrated UMAP, reflecting the local diversity of cell classes in each clustering neighborhood, based on embryonic origin, physiological function or tissue architecture. Data displayed as boxplot. Non-parametric Wilcoxon test.* ***G:*** *Integrated UMAP colored by cell identity defined during analysis on individual organ datasets.* ***H:*** *Percentage of BCs and SBCs present in each cluster presented in Figure S2B.* ***I:*** *Boxplots show the null distributions of mean η^2 values obtained by permutation of Tissue and Identity labels (grey), while large black points indicate the observed mean η^2 for each variance component. Data displayed as boxplot, Permutation p-values (See Methods).* ***J:*** *Module decomposition of the cornea suprabasal transcriptome into Invariant, Suprabasal-core, Squamous-core, Cornea-core and Cornea-suprabasal modules (See Methods).* ***K:*** *Schematic Russian-doll decomposition of single cell transcriptome.* ***L:*** *Expression-weighted variance decomposition highlights divergent influences on basal and suprabasal cells. Expression-weighted η^2 values were aggregated at the single-cell level to quantify the relative influence of identity, function, tissue, and interaction terms on transcriptional variance. Data displayed as boxplot. Non parametric Wilcoxon test.*

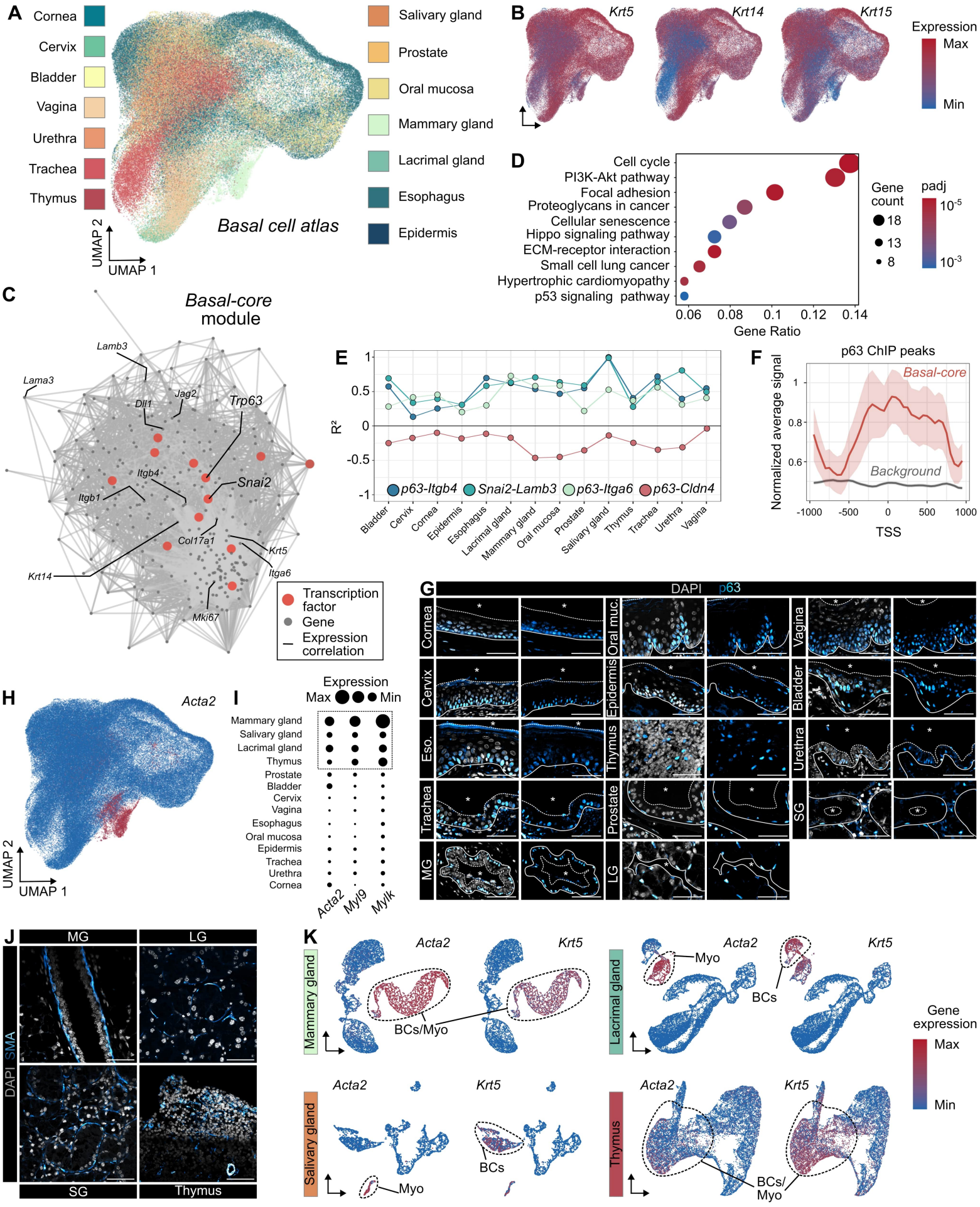

A
Cornea
Cervix
Bladder
Vagina
Urethra
Trachea
Thymus
Salivary gland
Prostate
Oral mucosa
Mammary gland
Lacrimal gland
Esophagus
Epidermis
Basal cell atlas
UMAP 2
UMAP 1
B
Krt5
Krt14
Krt15
Expression
Max
Min
C
Basal-core module
Lamb3
Lama3
Dll1
Jag2
Trp63
Itgb4
Snai2
Itgb1
Col17a1
Krt5
Krt14
Itga6
Mki67
Transcription factor
Gene
Expression correlation
D
Cell cycle
PI3K-Akt pathway
Focal adhesion
Proteoglycans in cancer
Cellular senescence
Hippo signaling pathway
ECM-receptor interaction
Small cell lung cancer
Hypertrophic cardiomyopathy
p53 signaling pathway
Gene Ratio
Gene count
18
13
8
padj
E
R²
p63-Itgb4
Snai2-Lamb3
p63-Itga6
p63-Cldn4
Bladder
Cervix
Cornea
Epidermis
Esophagus
Lacrimal gland
Mammary gland
Oral mucosa
Prostate
Salivary gland
Thymus
Trachea
Urethra
Vagina
F
p63 ChIP peaks
Basal-core
Background
Normalized average signal
TSS
G
DAPI
p63
Cornea
Cervix
Eso.
Trachea
MG
Oral muc.
Epidermis
Thymus
Prostate
LG
Vagina
Bladder
Urethra
SG
H
Acta2
UMAP 2
UMAP 1
I
Expression
Max
Min
Mammary gland
Salivary gland
Lacrimal gland
Thymus
Prostate
Bladder
Cervix
Vagina
Esophagus
Oral mucosa
Epidermis
Trachea
Urethra
Cornea
Acta2
Myl9
Mylk
J
MG
LG
DAPI SMA
SG
Thymus
K
Acta2
Krt5
Mammary gland
BCs/Myo
Lacrimal gland
Myo
BCs
Salivary gland
Thymus
BCs/Myo
Gene expression
Max
Min

Figure 3. Characterization of the basal compartment: similarities and specialization across tissues.

A: *Integrated UMAP of basal cells scRNAseq data colored by tissue.* ***B:*** *Integrated UMAP of scRNA-seq data colored by expression level of Krt5, Krt14 and Krt15 in individual cells.* ***C:*** *Basal-core GRN obtained by variance decomposition. Representation of the Basal-core as a correlation network of shared basal genes across tissues, highlighting transcription factors (red) as central hubs within the basal GRN.* ***D:*** *KEGG pathway enrichment across Basal-core module.* ***E:*** *Pairwise co-expression values for different selected correlating or anticorrelating gene pairs across the 14 tissues.* ***F****: Metagene profile of p63 ChIP-seq signal centered on transcription start sites (TSS) of basal-core genes (red) and background promoters (grey). The mean normalized signal (± standard error of the mean) is plotted in 50 bp bins across a ±1 kb window around the TSS. The p63 signal was capped at the 99.9th percentile and normalized for sequencing depth per chromosome.* ***G:*** *Confocal imaging of adult WT multilayered epithelia stained for Dapi and p63 (Cyan). *: indicate lumen or external space.* ***H:*** *Integrated UMAP of basal cells scRNAseq data colored by level of Acta2 gene expression.* ***I****: Dot plot showing percentage level of expression of genes associated with contractility (Acta2, Myl9 and Mylk) in the 14 murine multilayered epithelia.* ***J:*** *Confocal imaging of adult WT mouse section of Mammary, Lacrimal, Salivary gland and Thymus stained for Dapi (grey) and α-SMA (Cyan).* ***I:*** *UMAP of epithelial cells scRNA-seq data from Mammary, Salivary, Lacrimal gland and Thymus, colored by expression of Krt5 or Acta2 in individual cells. *: indicate lumen or external space; dotted line: indicate the interface between apical domain of epithelia and external space; continuous line: indicate the interface between epithelium and basement membrane. In G and J, scale bar: 50 µm.*

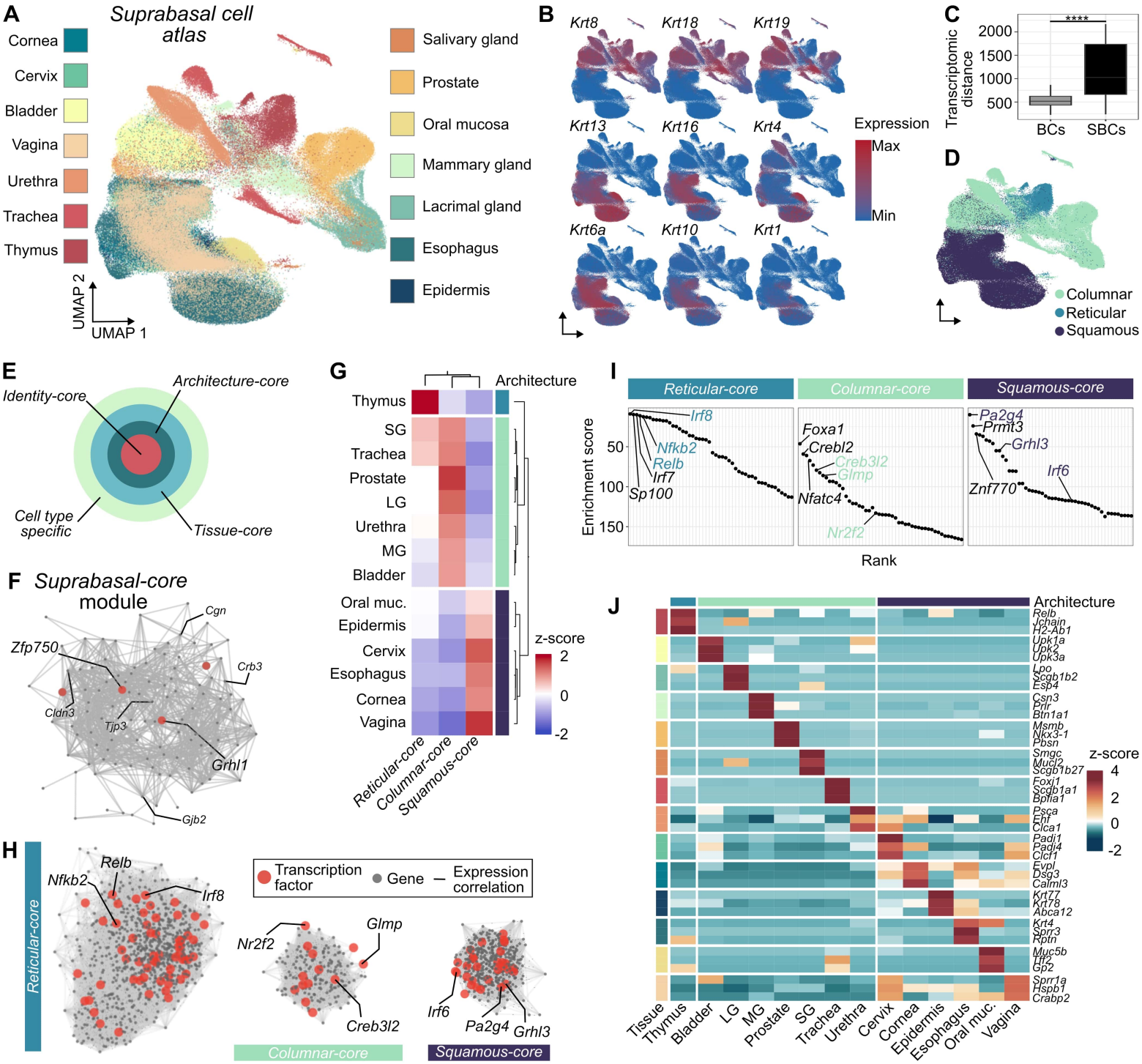
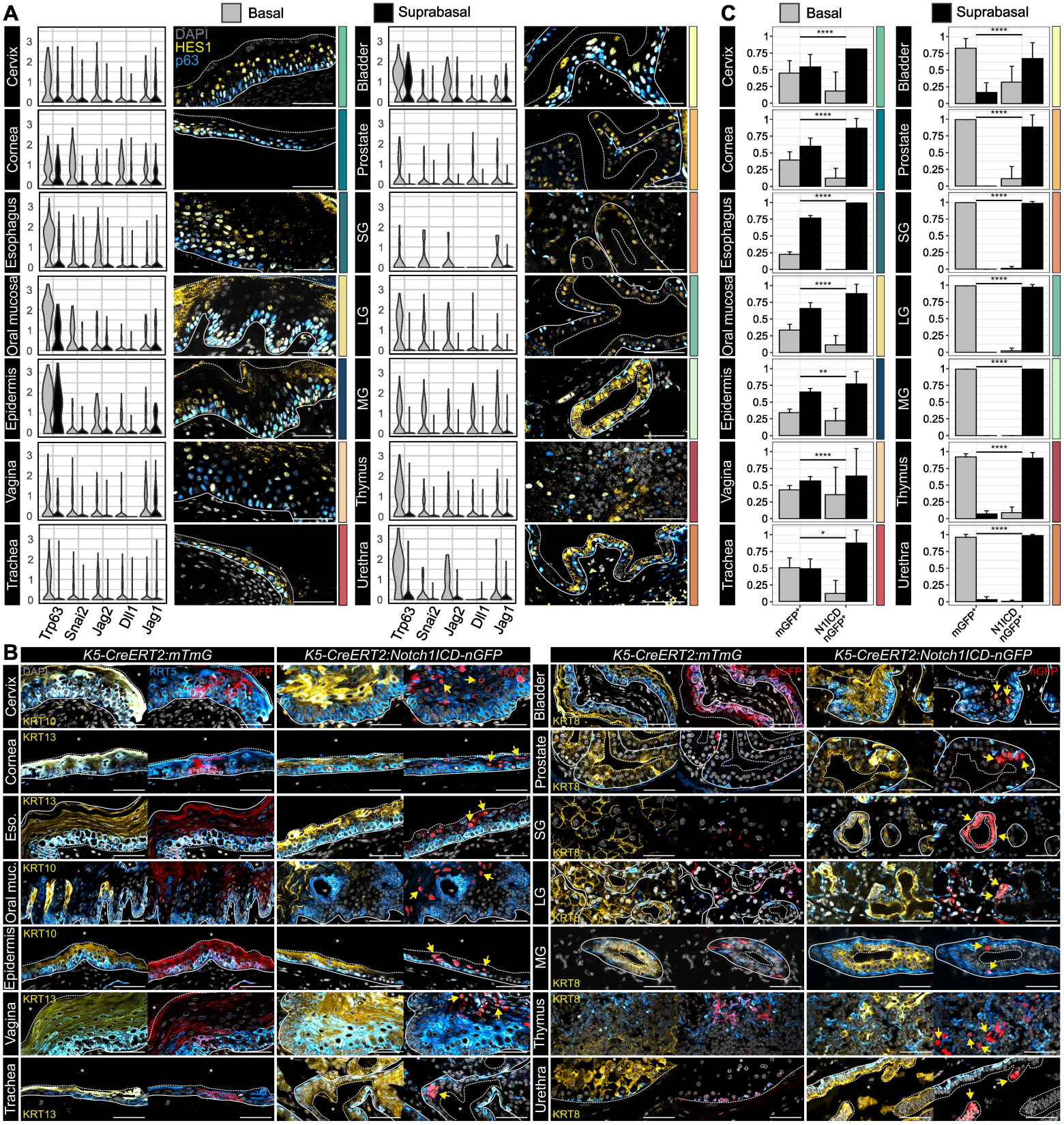
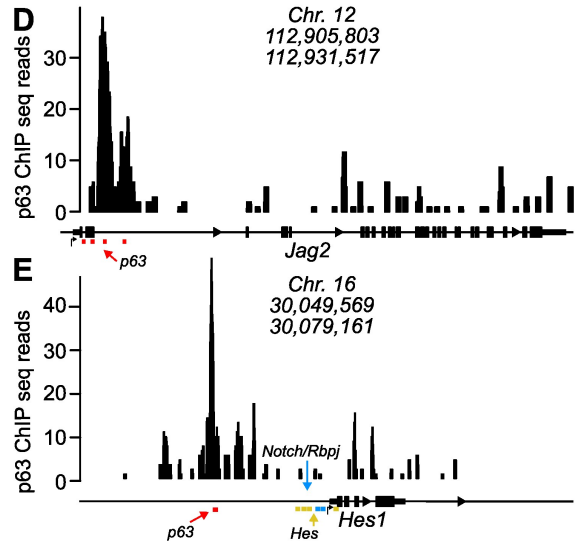

A
Suprabasal cell atlas
Cornea
Cervix
Bladder
Vagina
Urethra
Trachea
Thymus
Salivary gland
Prostate
Oral mucosa
Mammary gland
Lacrimal gland
Esophagus
Epidermis
UMAP 2
UMAP 1
B
Krt8
Krt18
Krt19
Krt13
Krt16
Krt4
Krt6a
Krt10
Krt1
Expression
Max
Min
C
Transcriptomic distance
2000
1500
1000
500
****
BCs
SBCs
D
Columnar
Reticular
Squamous
E
Identity-core
Architecture-core
Cell type specific
Tissue-core
F
Suprabasal-core module
Cgn
Zfp750
Crb3
Cldn3
Tjp3
Grhl1
Gjb2
G
Architecture
Thymus
SG
Trachea
Prostate
LG
Urethra
MG
Bladder
Oral muc.
Epidermis
Cervix
Esophagus
Cornea
Vagina
z-score
2
0
-2
Reticular-core
Columnar-core
Squamous-core
I
Reticular-core
Columnar-core
Squamous-core
Enrichment score
50
100
150
Irf8
Nfkb2
Relb
Irf7
Sp100
Foxa1
Creb3l2
Creb3l2
Glmp
Nfatc4
Nr2f2
Pa2g4
Prmt3
Grhl3
Irf6
Znf770
Rank
H
Reticular-core
Nfkb2
Relb
Irf8
Transcription factor
Gene
Expression correlation
Nr2f2
Glmp
Creb3l2
Columnar-core
Irf6
Pa2g4
Grhl3
Squamous-core
J
Architecture
Relb
Ighain
H2-Ab1
Upk1a
Upk2
Upk3a
Lpo
Scgb1b2
Esp4
Csn3
Prlr
Btn1a1
Msmb
Nkx3-1
Pbsn
Smgc
Muc12
Scgb1b27
Foxj1
Scgb1a1
Bpifa1
Psca
Ehf
Clca1
Padi1
Padi4
Clcf1
Evpl
Dsg3
Calml3
Krt77
Krt78
Abca12
Krt4
Sprr3
Krtn
Muc5b
Tff2
Gp2
Sprr1a
Hspb1
Crabp2
z-score
4
2
0
-2
Tissue
Thymus
Bladder
LG
MG
Prostate
SG
Trachea
Urethra
Cervix
Cornea
Epidermis
Esophagus
Oral muc.
Vagina

Figure 4. Diversification of the suprabasal transcriptome in relation to organ function

A: *Integrated UMAP of suprabasal cells scRNAseq data colored by tissue.* ***B:*** *Integrated UMAP of scRNA-seq data colored by expression level of keratins in individual cells.* ***C:*** *Euclidian distance across Basal and Suprabasal pseudo bulk samples. Data displayed as boxplot. Non parametric Wilcoxon test.* ***D:*** *Integrated UMAP of suprabasal cells scRNAseq data colored by tissue architecture (Columnar, Squamous, Reticular).* ***E:*** *Schematic Russian-doll decomposition of single cell transcriptome of suprabasal cells.* ***F:*** *Correlation network of the Suprabasal core module, highlighting transcription factors (red) as central hubs within the suprabasal GRN.* ***G:*** *Sparse regression–derived module weights are shown as a z-scored heatmap. Rows correspond to tissues, grouped by architectural class (Squamous, Columnar, Reticular), and columns correspond to architectural core modules. Color scale indicates relative module activity.* ***H:*** *Correlation network of the architecture core modules, highlighting transcription factors (red) as central hubs within the suprabasal GRNs of the indicated tissues.* ***I****: Transcription factor motif enrichment across architectural modules. Enrichment analysis of transcription factor–binding motifs performed with the ChEA database on genes defining the Columnar, Reticular, and Squamous core modules. Each plot shows the ranked enrichment scores of transcription factors, with representative top regulators annotated. Colored annotated TFs correspond to genes that belongs to the module.* ***J:*** *Z-score expression of Tissue and Suprabasal cell specific genes (3 genes per module).*


A
Basal
Suprabasal
DAPI
HES1
p63
Cervix
Cornea
Esophagus
Oral mucosa
Epidermis
Vagina
Trachea
Bladder
Prostate
SG
LG
MG
Thymus
Urethra
Trp63
Snai2
Jag2
Dll1
Jag1
C
mGFP+
N1ICD nGFP+
B
K5-CreERT2:mTmG
K5-CreERT2:Notch1ICD-nGFP
KRT10
KRT13
KRT8
Eso.
Oral muc.


D
Chr. 12
112,905,803
112,931,517
p63 ChIP seq reads
p63
Jag2
E
Chr. 16
30,049,569
30,079,161
Notch/Rbpj
Hes
Hes1

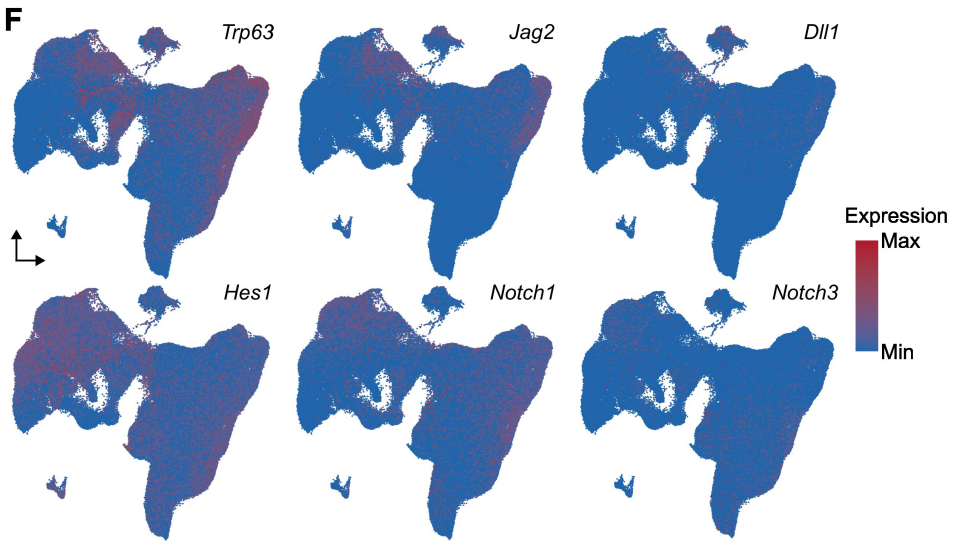

F
Trp63
Jag2
Dll1
Hes1
Notch1
Notch3
Expression
Max
Min

Figure 5. Notch signaling as a driver of suprabasal lineage commitment (Preliminary data)

A: *Gene expression of Trp63, Snai2 and Notch ligands Dll1, Jag1 and Jag2 across BCs and SBCs of all 14 epithelia and confocal imaging of adult WT multilayered epithelia stained for Dapi, p63 and HES1.* ***B:*** *Confocal imaging of immunostaining for K5, SBCs marker and membrane-bound GFP (mGFP) in K5-Cre^{ERT2}:mTmG (Control)or nuclear GFP (nGFP) in K5-Cre^{ERT2}:N1ICD-nGFP (Notch GOF) epithelia after lineage tracing.* ***C:*** *Quantification of fate of mGFP^{+} or nGFP^{+} cells after 7 weeks of lineage tracing for mGFP^{+},1 week (Cervix, Cornea, Esophagus, Oral mucosa, Epidermis, Vagina and Trachea) or 7 weeks for nGFP^{+} (Bladder, Prostate, Salivary gland (SG), Lacrimal gland (LG), Mammary gland (MG), Thymus and Urethra). Data are displayed as Mean + s.d. Chi-squared test (n=3 biological replicates).* ***D-E:*** *ΔNp63 ChIP-seq read coverage across Jag2 (D) and Hes1 (E) loci (mm10 assembly). Signal tracks represent aligned ChIP-seq raw coverage for ΔNp63 in mouse keratinocyte. The predicted p63-binding site are colored in red. The approximate positions of two Notch/RBP-Jκ sites (blue)*[1] *and four Hes-binding sites (yellow)*[2] *are also shown.* ***F:*** *Integrated UMAP of scRNA-seq data colored by expression level of Trp63, Jag2, Dll1, Hes1, Notch1 and Notch3 in individual cells. *: indicate lumen or external space ; dotted line: indicate the interface between apical domain of epithelia and external space; continuous line: indicate the interface between epithelium and basement membrane. Scale bar: 50 µm.*

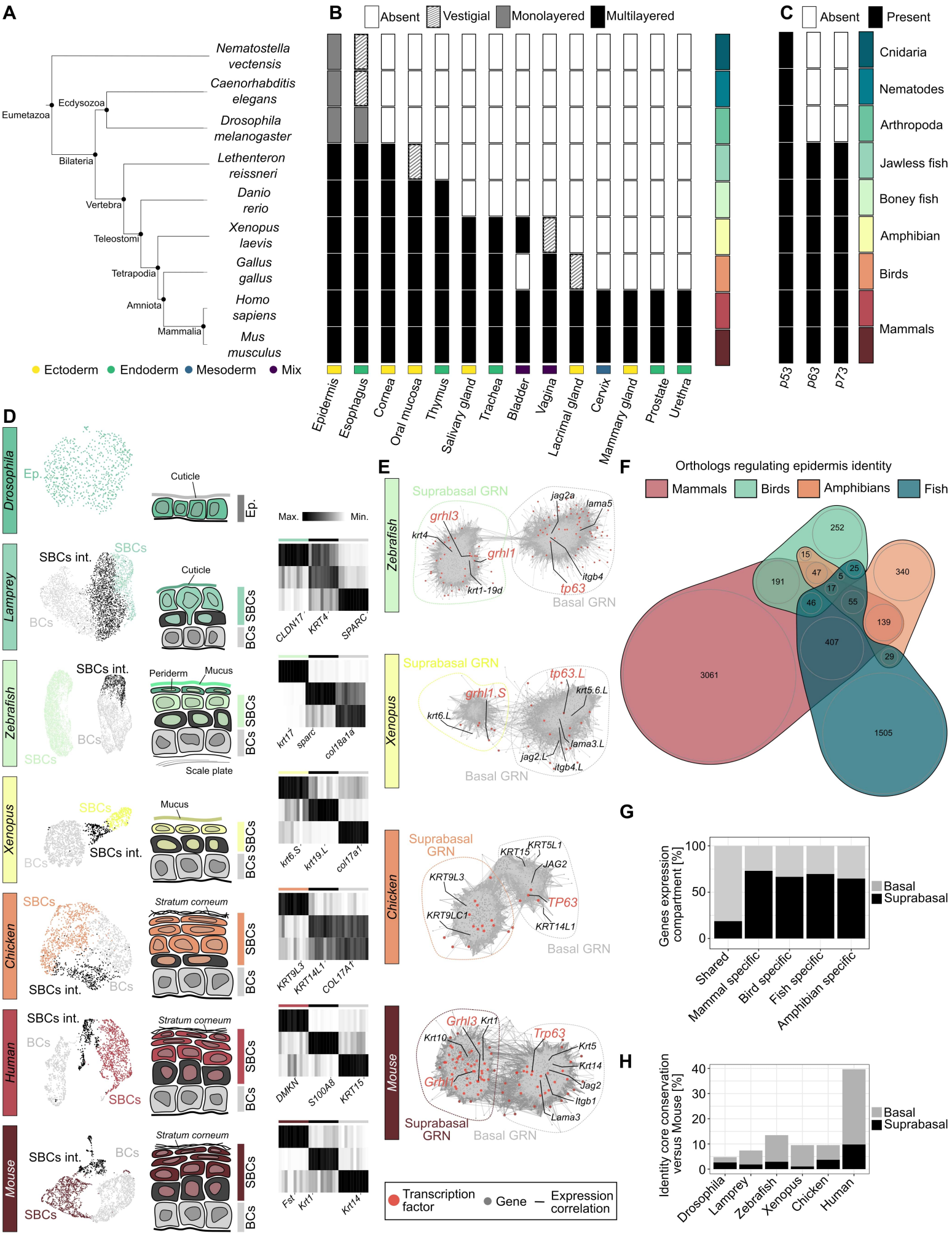

A
Nematostella vectensis
Caenorhabditis elegans
Drosophila melanogaster
Lethenteron reissneri
Danio rerio
Xenopus laevis
Gallus gallus
Homo sapiens
Mus musculus
Eumetazoa
Ecdysozoa
Bilateria
Vertebra
Teleostomi
Tetrapodia
Amniota
Mammalia
Ectoderm
Endoderm
Mesoderm
Mix
B
Absent
Vestigial
Monolayered
Multilayered
Epidermis
Esophagus
Cornea
Oral mucosa
Thymus
Salivary gland
Trachea
Bladder
Vagina
Lacrimal gland
Cervix
Mammary gland
Prostate
Urethra
C
Absent
Present
Cnidaria
Nematodes
Arthropoda
Jawless fish
Boney fish
Amphibian
Birds
Mammals
p53
p63
p73
D
Drosophila
Lamprey
Zebrafish
Xenopus
Chicken
Human
Mouse
Ep.
Cuticle
SBCs int.
SBCs
BCs
Periderm
Mucus
Scale plate
Stratum corneum
Max.
Min.
CLDN17
KRT4
SPARC
krt17
sparc
col18a1a
krt6.S
krt19.L
col17a1
KRT9L3
KRT14L1
COL17A1
DMKN
S100A8
KRT15
Fst
Krt1
Krt14
E
Suprabasal GRN
Basal GRN
grhl3
krt4
grhl1
krt1-19d
jag2a
lama5
itgb4
tp63
tp63.L
krt5.6.L
grhl1.S
krt6.L
lama3.L
jag2.L
itgb4.L
KRT5L1
KRT15
JAG2
KRT9L3
TP63
KRT9LC1
KRT14L1
Grhl3
Krt1
Krt10
Trp63
Krt5
Krt14
Grhl1
Jag2
Itgb1
Lama3
Transcription factor
Gene
Expression correlation
F
Orthologs regulating epidermis identity
Mammals
Birds
Amphibians
Fish
252
15
47
5
25
191
17
340
46
55
139
407
29
3061
1505
G
Genes expression compartment [%]
100
50
0
Basal
Suprabasal
Shared
Mammal specific
Bird specific
Fish specific
Amphibian specific
H
Identity core conservation versus Mouse [%]
40
30
20
10
0
Basal
Suprabasal
Drosophila
Lamprey
Zebrafish
Xenopus
Chicken
Human

Figure 6. Identification of an ancestral GRN regulating epithelial stratification

A-B: *Phylogenetic tree (A) and the associated presence or absence of multilayered epithelia within each organism (B). Tissues are color-coded by the embryonic germ layer of origin.* ***C:*** *Presence or absence of genes of the p53 family within each organism.* ***D:*** *UMAP of scRNA-seq data colored by organ and cell cluster across Drosophila, Lamprey, Zebrafish, Xenopus, Chicken, mouse and Human epidermis. Associated scheme of the epidermis structure and marker of each cell population. Ep.: Epithelium, int.: intermediate.* ***E:*** *Gene co-expression network in each epidermis scRNA seq dataset.* ***F:*** *Venn diagram of the murine orthologs present in the co-expression network across species in epidermis.* ***G:*** *Percentage of genes expressed in basal or suprabasal compartment of the epidermis for shared or species-specific group.* ***H:*** *Proportion of murine orthologs from Basal and Suprabasal-core modules detected in the epidermal GRN of each species.*

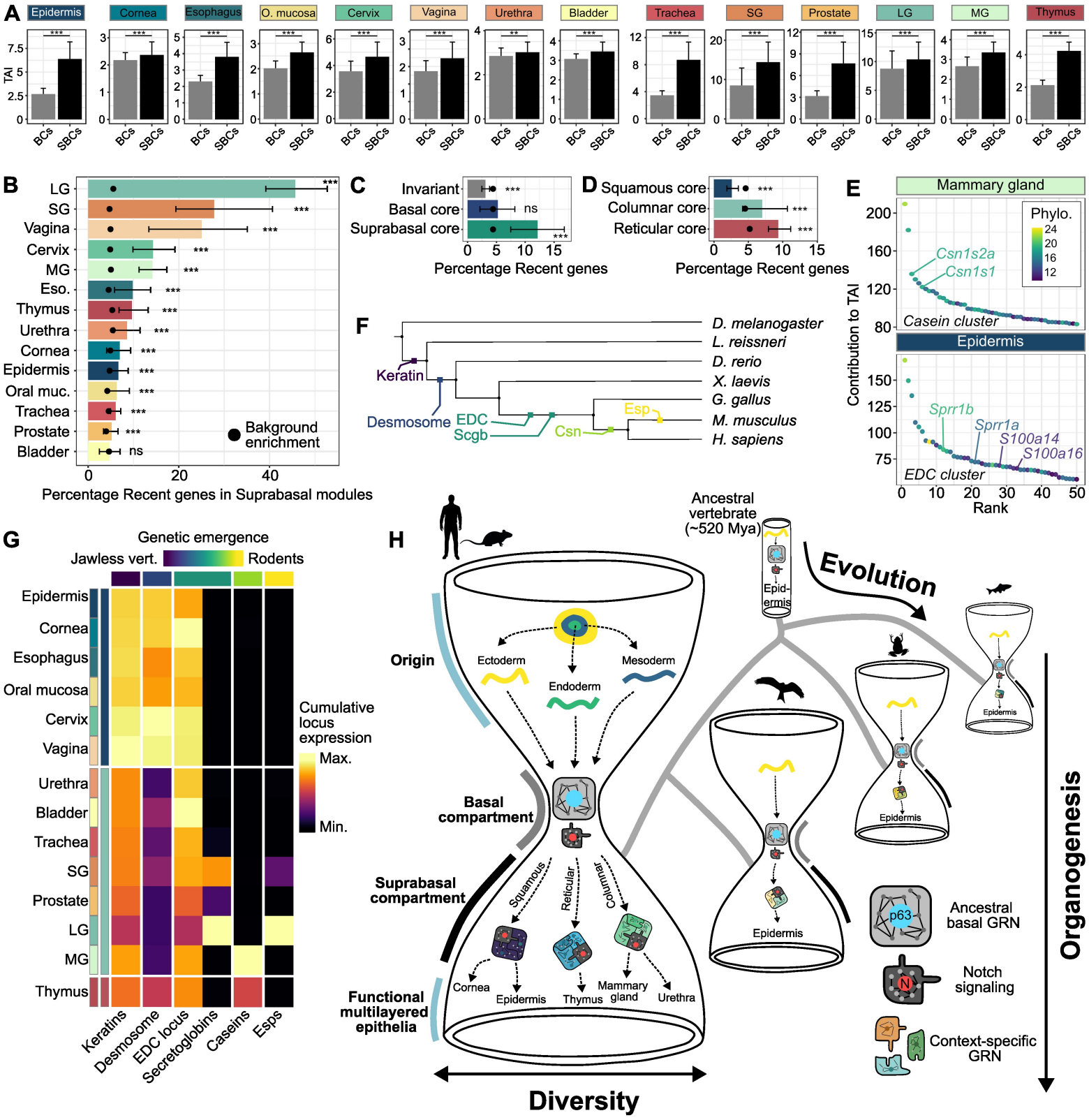

A
Epidermis
Cornea
Esophagus
O. mucosa
Cervix
Vagina
Urethra
Bladder
Trachea
SG
Prostate
LG
MG
Thymus
TAI
BCs
SBCs
B
LG
SG
Vagina
Cervix
MG
Eso.
Thymus
Urethra
Cornea
Epidermis
Oral muc.
Trachea
Prostate
Bladder
Bakground enrichment
Percentage Recent genes in Suprabasal modules
C
Invariant
Basal core
Suprabasal core
Percentage Recent genes
D
Squamous core
Columnar core
Reticular core
Percentage Recent genes
E
Mammary gland
Phylo.
Csn1s2a
Csn1s1
Casein cluster
Epidermis
Sprr1b
Sprr1a
S100a14
S100a16
EDC cluster
Contribution to TAI
Rank
F
D. melanogaster
L. reissneri
D. rerio
X. laevis
G. gallus
M. musculus
H. sapiens
Keratin
Desmosome
EDC
Scgb
Csn
Esp
G
Genetic emergence
Jawless vert.
Rodents
Epidermis
Cornea
Esophagus
Oral mucosa
Cervix
Vagina
Urethra
Bladder
Trachea
SG
Prostate
LG
MG
Thymus
Cumulative locus expression
Max.
Min.
Keratins
Desmosome
EDC locus
Secretoglobins
Caseins
Esps
H
Ancestral vertebrate (~520 Mya)
Epid-ermis
Evolution
Origin
Ectoderm
Endoderm
Mesoderm
Basal compartment
Suprabasal compartment
Squamous
Reticular
Columnar
Functional multilayered epithelia
Cornea
Epidermis
Thymus
Mammary gland
Urethra
Epidermis
Epidermis
Epidermis
p63
Ancestral basal GRN
Notch signaling
Context-specific GRN
Organogenesis
Diversity

Figure 7. Evolution of tissue specific transcriptome through gene toolkit evolution and GRN rewiring

A: *Bootstrapped TAI calculated in pseudo bulk basal and suprabasal samples based on identity specific genes. Data displayed as mean + sd. Non parametric Wilcoxon test.* ***B****: Bootstrapped Enrichment in Recent genes (phylostratum >8) in tissue specific suprabasal modules compared to background enrichment in the same sample. Data displayed as mean + sd. Non parametric Wilcoxon test.* ***C****: Bootstrapped Enrichment in Recent genes in invariant and identity core modules compared to background enrichment in whole genome. Data displayed as mean + sd. Non parametric Wilcoxon test.* ***D****: Bootstrapped Enrichment in Recent genes in architecture specific modules compared to background enrichment in related sample. Data displayed as mean + sd. Non parametric Wilcoxon test.* ***E:*** *Main contributor to TAI in Mammary gland and Epidermis. Genes from casein and EDC complex are annotated respectively.* ***F:*** *Phylogenetic tree and associated genetic event related to emergence of each gene cluster.* ***G:*** *Cumulative expression of gene clusters in each murine multilayered epithelium.* ***H.*** *Developmental and evolutionary hourglass model of multilayered epithelia. During gastrulation, epithelia of distinct germ-layer origins (ectoderm, mesoderm, endoderm) emerge (top). Across vertebrates, these diverse embryonic inputs converge on a conserved basal bottleneck governed by the p63–Notch regulatory module (waist), representing a developmental and evolutionary constraint inherited from the ancestral vertebrate (~520 Mya). From this shared basal program, lineage- and tissue-specific suprabasal GRNs (bottom) drive the diversification of multilayered epithelia into squamous, reticular, and columnar architectures.*