Virology 537 (2019) 14–19

Contents lists available at ScienceDirect

Virology

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

Single-cell RNA sequencing reveals cell type-specific HPV expression in hyperplastic skin lesions

Katharina Devitt[a], Sarah J. Hanson[a], Zewen K. Tuong[a], Erin McMeniman[b,c], H. Peter Soyer[b,c], Ian H. Frazer[a,*,1], Samuel W. Lukowski[d,1]

[a] The University of Queensland Diamantina Institute, The University of Queensland, Woolloongabba, QLD, 4102, Australia
[b] The University of Queensland Diamantina Institute, The University of Queensland, Dermatology Research Centre, Woolloongabba, QLD, 4102, Australia
[c] Department of Dermatology, The Princess Alexandra Hospital, Woolloongabba, QLD, 4102, Australia
[d] Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD, 4072, Australia

ARTICLE INFO

Keywords:
HPV
Single-cell RNA-seq
Transcription
Immunosuppression
Epithelial

ABSTRACT

Human Papillomavirus infection is highly prevalent worldwide. While most types of HPV cause benign warts, some high-risk types are known to cause cervical cancer, as well as cancer of the oral cavity and head and neck. Persistent cutaneous HPV infection can be particularly problematic in patients with chronic immunosuppression, for example following organ transplantation. Due to unknown mechanisms, these patients may develop numerous warts, as well as present with a dramatically increased skin cancer prevalence.

Despite an association between HPV persistence in the epidermis and excessive wart or squamous cancer development, the molecular mechanisms linking immunosuppression, HPV expression and excessive epidermal proliferation have not been determined, largely due to low-sensitivity methodology to capture rare viral transcription events. Here, we use single-cell RNA sequencing to profile HPV-positive skin lesions from an immunosuppressed patient that were found to express the alphapapillomavirus HPV78 in basal keratinocytes, suprabasal keratinocytes and hair follicle stem cells. This method can be applied to detect and investigate HPV transcripts in cutaneous lesions, allowing mechanistic links between immunosuppression-induced HPV life cycle and epidermal hyperproliferation to be uncovered.

1. Introduction

The causality between persistent α-HPV infection of epithelial cells and cancer of the cervix and head and neck has been well established. Long-term immunosuppressive treatment, such as is necessary after organ transplant, leads to dramatically increased incidences of warts, anogenital neoplasias as well as squamous skin cancers (Larsen et al., 2019; Ozsaran et al., 1999). While occurrences of these lesions following organ transplantation have been suggested to stem from the reactivation of latent HPV (Maglennon et al., 2014), this has not yet been shown conclusively. Animal models allow for studies of related viruses (such as Cottontail Rabbit Papillomavirus, CRPV), but not HPV. In humans, the study of viral reactivation is challenging and has in the past been difficult due to a lack of sensitive enough methods to detect rare viral transcription in single cells. Only recently, methods have become available to detect rare transcripts with high sensitivity at single-cell level. We have previously shown that single-cell RNA sequencing (scRNA-seq) approaches can be used to identify HPV16 *E7* transcripts in epidermal keratinocytes (Lukowski et al., 2018). Here, we present a streamlined workflow (Fig. 1) that incorporates HPV genotyping in epidermis prior to single-cell RNA sequencing via the 10X Genomics Chromium platform (10X Genomics, Pleasanton, CA). Using three warts from an immunosuppressed organ recipient, we identified transcription of two HPV types in hyperproliferative epidermal lesions. Here, we show that scRNA-seq is a powerful tool to interrogate the viral life cycle, pathways and individual target molecules in the context of viral reactivation-induced hyperplasia.

2. Materials and methods

2.1. Collection of human samples

Three warts (common warts by histopathology: two from the chest, one from the elbow) and one normal skin (inner arm, not sun-exposed)

* Corresponding author.
E-mail address: i.frazer@uq.edu.au (I.H. Frazer).
[1] Senior author.

https://doi.org/10.1016/j.virol.2019.08.007
Received 29 May 2019; Received in revised form 8 August 2019; Accepted 8 August 2019
Available online 13 August 2019

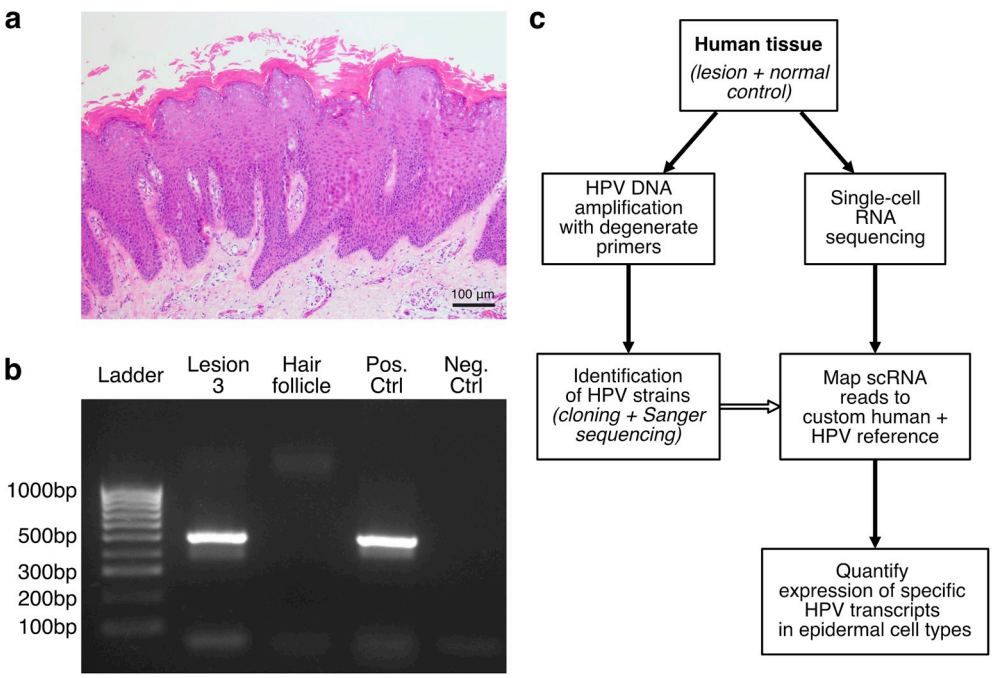

Fig. 1. Pipeline for detection of HPV transcription in human epidermis. (a) Histological section of a representative warty lesion stained with Haematoxylin and Eosin. Scale bar = 100 µm. Three warts and one normal skin sample from a patient with multiple warts were collected. DNA from lesion 3 (elbow) was genotyped with broad-spectrum HPV-specific primer sets (b). All four samples were processed into single-cell suspensions, cleared of dead cells, doublets and debris, and sequenced by 3′ scRNA-seq, using the workflow shown in (c). The sequences of the identified HPV types were combined with the human reference for scRNA-seq read mapping (empty arrow). Pos ctrl: Caski cell DNA; neg ctrl: ultrapure dH_2O.

punch biopsy were sampled from a 38-year old immunosuppressed patient who presented to the Dermatology Department at the Princess Alexandra Hospital in 2018 with multiple warts in sun-exposed areas. The patient was chronically immunosuppressed after kidney transplantations in 2000 and 2017 and presented with numerous non-malignant warts in sun-exposed areas (chest, forehead, lower legs). The patient was advised on the research protocol, and written informed consent was obtained. The study was approved by The University of Queensland Human Research Ethics Committee (HREC-11-QPAH-477, The University of Queensland clearance no. 2012000052).

2.2. Genotyping by DNA PCR and cloning

A small piece of tissue from one sample, and a hair follicle from the normal skin was used for DNA extraction and HPV genotyping using the FAP59/64 degenerate primers that detect a broad range of common epithelial HPV types (Forslund et al., 2007; Li et al., 2013). The tissue was digested in lysis buffer (REDExtract-N-Amp™ PCR ReadyMix™, Merck, Darmstadt, Germany) for 30 min. PCR products were run on 1.5% agarose gel and FAP-amplified bands were excised. PCR products were cloned into competent *E. coli* using pGEM-T vectors as per the manufacturer's guidelines (Promega, Wisconsin, USA) and each cloned product was sequenced using the Sanger method by the Australian Genome Research Facility (AGRF). Sequencing of the cloned PCR products showed specific amplification of DNA matching two α-HPV genotypes with 98–99% sequence identity, HPV78 and HPV3, identified using BLAST.

2.3. Isolation of single cells from patient lesions

The remaining tissue was processed to obtain single-cell suspensions for scRNA-seq. The tissue was placed into 4 mg/mL Dispase overnight at 4 °C. The next day, the epidermis was peeled off and incubated in 0.25% Trypsin for 2 min at 37 °C. The reaction was stopped with serum-containing medium and the cells were washed through a 70 µm cell strainer, followed by two 40 µm filtrations. Fluorescence-activated cell sorting resulted in single-cell suspensions of 8.2×10^4–1.3×10^5 cells with > 77% viability from the lesion and 1.2×10^4 cells with a viability of 80% from normal tissue.

2.4. scRNA-seq

scRNA-seq was performed on the normal and lesion samples using the 10X Genomics Chromium 3′ Gene Expression assay (Pleasanton, CA). The Chromium instrument was used to partition viable cells with barcoded beads, and cDNA from each cell was prepared using the Single Cell 3′ Library, Gel Bead and Multiplex Kit (version 2, PN-120234; 10X Genomics) per the manufacturer's instructions. Cell numbers in each reaction were optimized to capture approximately 10,000 cells and the libraries were pooled and sequenced on an Illumina NextSeq500, using a 150-cycle High Output reagent kit (NextSeq500/550 version 2, FC-404-2002; Illumina) in standalone mode as follows: 26 cycles (read 1), 8 cycles (I7 index), and 98 cycles (read 2). scRNA-seq was performed by the IMB Sequencing Core Facility.

2.5. Bioinformatics processing

Single-cell data was processed from raw Illumina BCL files with the *cellranger* pipeline, version 2.1 (*mkfastq*, *count*, *aggr*) (Zheng et al., 2017) using the default parameters. Reads were aligned to a custom reference genome comprising the human GRCh38 and the HPV 3 and HPV 78 genomes (Van Doorslaer et al., 2017) using the STAR aligner (Dobin et al., 2013) included in the *cellranger* pipeline. Cell barcode and unique molecular identifier (UMI) quality control was performed by *cellranger* during the *count* stage using default parameters. A between-sample normalized gene expression matrix for four samples was generated using *cellranger aggr* for further analysis.

2.6. scRNA-seq analysis

The aggregated single-cell gene expression data generated by *cellranger* was used as the input for the Seurat analysis software (version 2.3.4, Butler et al., 2018). Expression levels for each transcript were determined using the number of UMIs per transcript. QC and filtering steps were performed in Seurat to remove outlier genes and cells, such that genes expressed in three or more cells, and cells with 100–5000 genes were retained. Cells expressing greater than 10% mitochondrial genes were removed. Expression data was normalized using the Log-Normalize function, and the ScaleData function was used to regress out the variation driven by the number of UMIs and mitochondrial gene

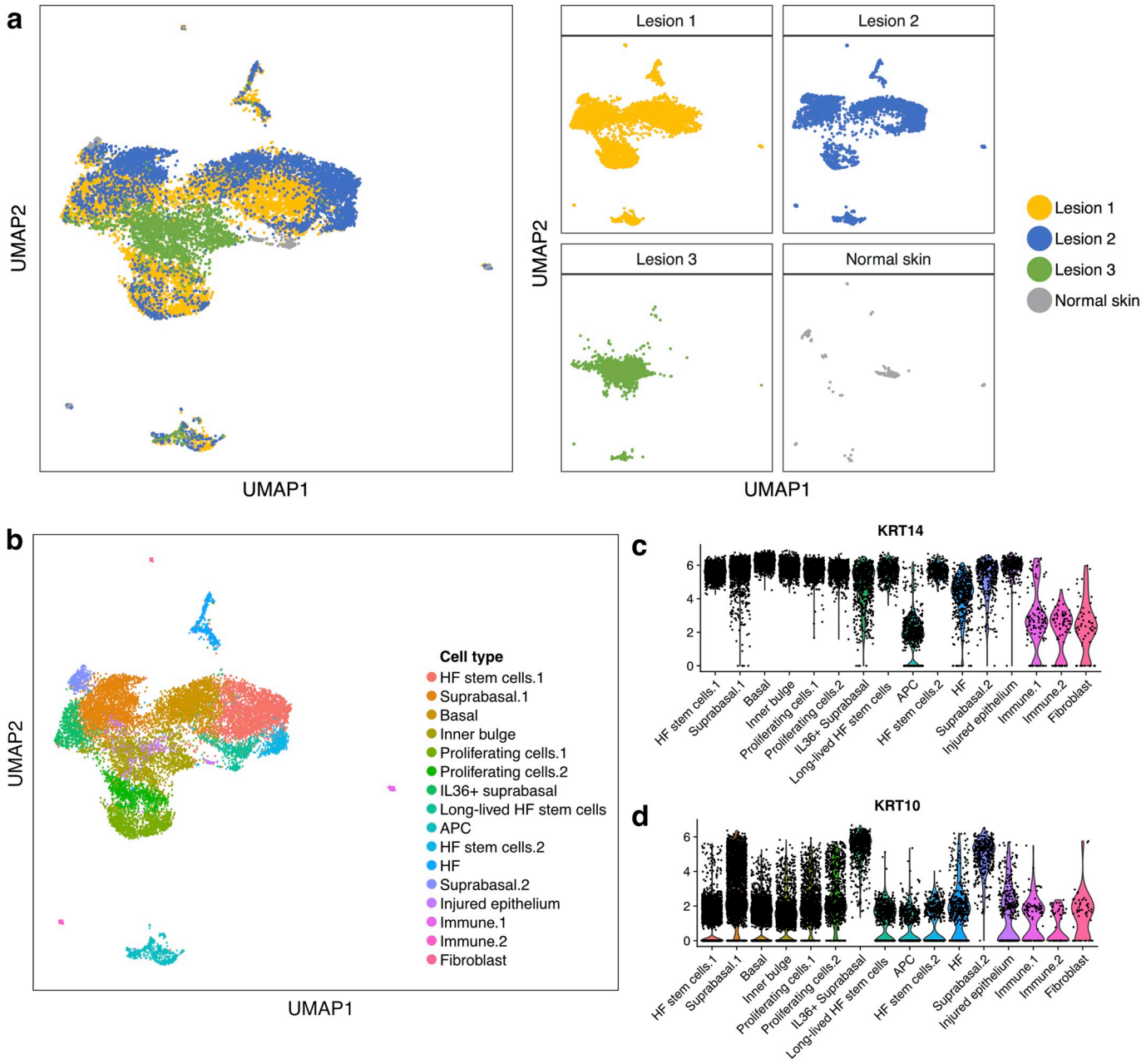

Fig. 2. Single-cell RNA sequencing reveals distinct populations between the lesions. (a) UMAP plot showing the cells from three warts and one normal skin sample. (b) UMAP plot of cell clusters and their cell type as classified by their expression profiles. (c) Single cell expression of basal keratin marker *KRT14* and (d) the suprabasal marker *KRT10*. Violin plots show log-transformed expression of individual genes. Black dots represent individual cells.

Table 1
Proportion of cell type per sample. Absolute numbers of each cell type per lesion are given, as well as percentage proportion of each cell type.

	HF stem cells 1	Suprabasal 1	Basal	Inner bulge	Proliferating cells 1	Proliferating cells 2	IL36+ Suprabasal	Long-lived HF stem cells
Lesion 1	1342 (49%)	1386 (55%)	1289 (67%)	38 (3%)	976 (81%)	811 (87%)	497 (74%)	276 (53%)
Lesion 2	1393 (51%)	1111 (44%)	616 (32%)	12 (1%)	208 (17%)	86 (9%)	143 (21%)	165 (31%)
Lesion 3	3 (0%)	24 (1%)	17 (1%)	1453 (97%)	27 (2%)	28 (3%)	31 (5%)	2 (0%)
Normal skin	0 (0%)	1 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0%)	0 (0%)	82 (16%)

	APC	**HF stem cells 2**	**HF**	**Suprabasal 2**	**Injured epithelium**	**Immune 1**	**Immune 2**	**Fibroblast**
Lesion 1	222 (43%)	82 (18%)	287 (63%)	167 (51%)	17 (5%)	31 (28%)	3 (4%)	22 (42%)
Lesion 2	234 (45%)	382 (82%)	159 (35%)	125 (38%)	43 (13%)	68 (62%)	56 (77%)	25 (48%)
Lesion 3	53 (10%)	3 (1%)	9 (2%)	3 (1%)	229 (70%)	1 (1%)	5 (7%)	0 (0%)
Normal skin	7 (1%)	1 (0%)	4 (1%)	32 (10%)	36 (11%)	10 (9%)	9 (12%)	5 (10%)

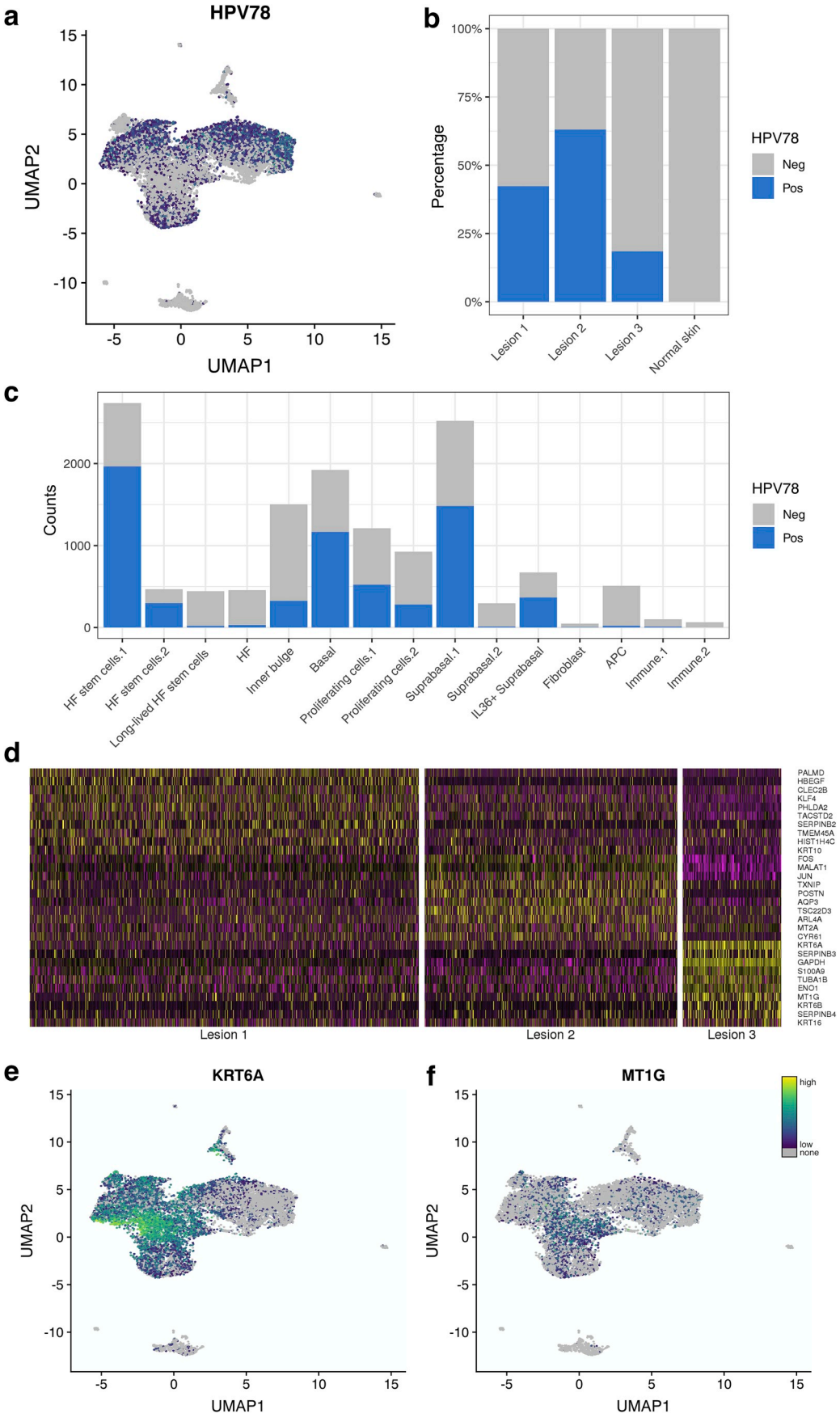

Fig. 3. HPV expression is unevenly distributed across the lesion samples. (a + b) HPV78 expression per cell overlaid on UMAP plot of all cells. Bar plot showing the proportion of cells per sample expressing HPV. (c) Bar plot showing the number of cells per cell type expressing HPV. (d) Heatmap of top 10 differentially expressed genes for each sample compared to combined remaining samples. Lesion 3 showed (e) high abundance of the stress-keratin 6 (*KRT6A*) and (f) increased expression of Metallothionein 1G (*MT1G*), a tumour suppressor that has been linked to HPV expression. The color gradients indicate gene expression levels from low (purple) to high (yellow).

expression. The first 20 principal components that explained the majority of variance in the data were used for clustering. To visualize the patterns of gene expression in each cell, Uniform Manifold Approximation and Projections (UMAP) (Becht et al., 2018) were generated using the principal component analysis-reduced data.

3. Results and discussion

Hyperplastic lesions sampled from an immunosuppressed patient were confirmed by pathology as HPV-positive common warts of plaque-like, flat appearance (Fig. 1). As the warts were HPV positive, they were deemed suitable candidates for droplet-based single cell RNA-sequencing and subsequent analysis, with the goal of using them as positive controls to establish the detection and analysis of HPV transcripts in human cutaneous lesions.

From each of the three proliferative lesion samples and the normal skin sample, we generated transcriptome data for 7,495, 5,017, and 1,951 cells from lesions 1–3 respectively, and 642 cells from normal skin. We obtained an average read depth of 1.0×10^5 (normal) and 2.8×10^4 (lesions) per cell, and an average of 2164 genes per cell in the lesion samples. Graph-based clustering, performed in Seurat (Butler et al., 2018), was used to group cells based on transcriptome similarity, and revealed 16 clusters that could be classified as discrete cell types (Fig. 2b). These were classified based on known marker genes determined by differential gene expression between clusters (Supplementary Table 1) and previously identified epidermal cell types defined by scRNA-seq and RNA-FISH (Joost et al., 2016). Each cell population clearly expressed dominant cell type marker genes (Fig. S2). We also observed Keratin 14 (*KRT14*; Fig. 2c), and Keratin 10 (*KRT10*; Fig. 2d) expression in almost all cell types, excluding the immune and fibroblast cell populations. The distributions of cells per sample and cluster are presented in Table 1.

Despite the high sequencing depth for the normal skin biopsy, we only detected expression from 276 genes. We attribute this to the high proportion of post-mitotic, Keratin 10-expressing (*KRT10*) cells present in the normal skin (Fig. 2a and d).

In addition to classifying the cell types present in the lesions, we also investigated the transcriptional differences between populations expressing HPV transcripts with those not expressing HPV. HPV expression was spread throughout all three lesions, with lesion 3 expressing considerably less HPV (Fig. 3a). We found that 45.12%, 63.07% and 19.97% of cells, in lesions 1–3 respectively, expressed at least one HPV transcript irrespective of strain, and of these, greater than 92% specifically expressed HPV78 transcripts in each lesion (Fig. 3b), with the remaining cells expressing HPV3, a strain closely related to HPV78 (Supplementary Table 2). No cells were found to express both HPV types concomitantly in these samples. However, the sensitivity and high resolution of the 10X Genomics Chromium technology allows for the detection and distinction of concomitantly expressed viral transcripts, if present. This is especially suited to viral transcripts with lower sequence identity than those from the HPV3 and HPV78 types detected here.

Analysis of each sample revealed heterogeneous marker gene expression in lesions 1 and 2 (Fig. 3d). In particular, lesion 1 and 2 showed upregulation of markers of altered skin barrier function (*ARL4A*, *MT2A*) and of inflammation, including the AP1 transcription factors *FOS* and *JUN* which are known to be upregulated during keratinocyte differentiation (Eckert et al., 2013; Mehic et al., 2005). Furthermore, lesion 2 showed increased expression of *FOXP1*, which is involved in hair follicle stem cell quiescence (Leishman et al., 2013) and was recently shown to play a role in melanoma progression (Donizy et al., 2018).

In lesion 3, we observed a high abundance of hair follicle (HF) inner bulge cells (Supplementary Table 3), increased expression of Keratin 6a (*KRT6A*) (Fig. 3e), and localised expression of metallothionein 1G (*MT1G*) (Fig. 3f), compared to lesions 1 and 2. MT1G is a known tumor suppressor and cancer biomarker that is methylated in cervical cancer but is upregulated in precancerous productive HPV16 infections (Henrique et al., 2005; Kang et al., 2018; Li et al., 2015). Our group has previously shown that expression of stress keratins such as Keratin 6 can modify the local immune environment and influence viral expression (Zhussupbekova et al., 2016). Furthermore, cell cycle analysis performed on each lesion revealed that lesion 3 contained a greater proportion of cells in G1 and fewer cells in S phase compared to lesions 1 and 2 (Fig. S1). Coupled with increased *MT1G* gene expression, the cells in lesion 3 may represent precancerous cells in the productive HPV lifecycle.

Our approach of combining HPV genotyping with scRNA-seq can be applied to detect HPV transcripts in cell preparations from skin tissue of different origin. It should be noted that the sequencing data from scRNA-seq 3′ gene expression assays is limited to the region around the poly-A site. Due to the polycistronic nature of HPV mRNAs, this could lead to a sequencing bias towards polyadenylated transcripts near the early and late poly-A sites. We were able to detect transcripts of all of the HPV genes, albeit with differing frequency. Since all early HPV transcripts are polycistronic and terminate immediately after the E5 gene, this detection method may lead to an overrepresentation of this gene. Furthermore, annealing of the oligo-dT at A-rich sites throughout the HPV genome is likely to generate reads at regions distal to the poly-A. Therefore, no conclusions were drawn about the abundance of individual transcripts, or their potential mechanisms, and HPV reads were considered collectively. The approach outlined here is suitable for the detection of rare transcriptional events at single-cell resolution. This has potentially significant implications for the detection of HPV transcripts as drivers of epithelial hyperplasia and malignancy and for subsequent development of a vaccine or immunotherapy treatment for one of the most abundant organ transplant complications.

Conflicts of interest

The authors declare no conflict of interest.

Funding sources

This work was supported by an NHMRC program grant (APP1071822) and by the Merchant Foundation.

This research was carried out at the Translational Research Institute, Woolloongabba, QLD 4102, Australia. The Translational Research Institute is supported by a grant from the Australian Government.

Acknowledgments

The authors thank the patient for donating her tissue for our research, Xiaohua (Christine) Shen for assistance with sample collection, Ryan Galea for cloning assistance, the UQ Diamantina Institute Flow Cytometry Core facility and the Institute for Molecular Bioscience Sequencing facility for their assistance with sample processing and sequencing and Joseph Powell and Annika Antonsson for valuable discussions. We also thank Dr. Rohan Mortimore, Medlab Pathology, for providing the photomicrograph.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2019.08.007.

Author contributions

KD and SWL contributed conceptualization, methodology, validation, analysis, investigation, writing (draft preparation, review, editing), visualisation, supervision, project administration; SWL was

responsible for software and data curation; SJH contributed to investigation; ZKT was involved in conceptualisation, software, formal analysis; EM contributed the clinical case, performed the procedure and provided input into the clinical analysis; HPS and IHF were involved in conceptualisation, supervision, project administration; IHF provided funding.

References

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., et al., 2018. Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. https://doi.org/10.1038/nbt.4314.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., 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.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Donizy, P., Pagacz, K., Marczuk, J., Fendler, W., Maciejczyk, A., Halon, A., et al., 2018. Upregulation of FOXP1 is a new independent unfavorable prognosticator and a specific predictor of lymphatic dissemination in cutaneous melanoma patients. OncoTargets Ther. https://doi.org/10.2147/OTT.S151286.

Eckert, R.L., Adhikary, G., Young, C.A., Jans, R., Crish, J.F., Xu, W., et al., 2013. AP1 transcription factors in epidermal differentiation and skin cancer. J. Skin Cancer. https://doi.org/10.1155/2013/537028.

Forslund, O., Iftner, T., Andersson, K., Lindelöf, B., Hradil, E., Nordin, P., et al., 2007. Cutaneous human papillomaviruses found in sun-exposed skin: beta-papillomavirus species 2 predominates in squamous cell carcinoma. J. Infect. Dis. 196, 876–883. https://doi.org/10.1086/521031.

Henrique, R., Jerónimo, C., Hoque, M.O., Nomoto, S., Carvalho, A.L., Costa, V.L., et al., 2005. MT1G hypermethylation is associated with higher tumor stage in prostate cancer. Cancer Epidemiol. Prev. Biomark. 14, 1274–1278. https://doi.org/10.1158/1055-9965.EPI-04-0659.

Joost, S., Zeisel, A., Jacob, T., Sun, X., La Manno, G., Lönnerberg, P., et al., 2016. Single-cell transcriptomics reveals that differentiation and spatial signatures shape epidermal and hair follicle heterogeneity. Cell Syst. 3, 221–237. https://doi.org/10.1016/j.cels.2016.08.010. e9.

Kang, S.D., Chatterjee, S., Alam, S., Salzberg, A.C., Milici, J., Burg van der, S.H., et al., 2018. Effect of productive human papillomavirus 16 infection on global gene expression in cervical epithelium. J. Virol. 92https://doi.org/10.1128/JVI.01261-18. e01261-18.

Larsen, H.K., Thomsen, L.T., Haedersdal, M., Dehlendorff, C., Sørensen, S.S., Kjaer, S.K., 2019. Risk of genital warts in renal transplant recipients—a registry-based, prospective cohort study. Am. J. Transplant. 19, 156–165. https://doi.org/10.1111/ajt.15056.

Leishman, E., Howard, J.M., Garcia, G.E., Miao, Q., Ku, A.T., Dekker, J.D., et al., 2013. Foxp1 maintains hair follicle stem cell quiescence through regulation of Fgf18. Development 140, 3809–3818. https://doi.org/10.1242/dev.097477.

Li, J., Pan, Y., Xu, Z., Wang, Q., Hang, D., Shen, N., et al., 2013. Improved detection of human papillomavirus harbored in healthy skin with FAP6085/64 primers. J. Virol. Methods 193, 633–638. https://doi.org/10.1016/j.jviromet.2013.06.026.

Li, L., Xu, C., Long, J., Shen, D., Zhou, W., Zhou, Q., et al., 2015. E6 and E7 gene silencing results in decreased methylation of tumor suppressor genes and induces phenotype transformation of human cervical carcinoma cell lines. Oncotarget 6, 23930–23943.

Lukowski, S.W., Tuong, Z.K., Noske, K., Senabouth, A., Nguyen, Q.H., Andersen, S.B., et al., 2018. Detection of HPV E7 transcription at single-cell resolution in epidermis. J. Investig. Dermatol. 138, 2558–2567. https://doi.org/10.1016/j.jid.2018.06.169.

Maglennon, G.A., McIntosh, P.B., Doorbar, J., 2014. Immunosuppression facilitates the reactivation of latent papillomavirus infections. J. Virol. 88, 710–716. https://doi.org/10.1128/JVI.02589-13.

Mehic, D., Bakiri, L., Ghannadan, M., Wagner, E.F., Tschachler, E., 2005. Fos and jun proteins are specifically expressed during differentiation of human keratinocytes. J. Investig. Dermatol. 124, 212–220. https://doi.org/10.1111/j.0022-202X.2004.23558.x.

Ozsaran, A.A., Ateş, T., Dikmen, Y., Zeytinoglu, A., Terek, C., Erhan, Y., et al., 1999. Evaluation of the risk of cervical intraepithelial neoplasia and human papilloma virus infection in renal transplant patients receiving immunosuppressive therapy. Eur. J. Gynaecol. Oncol. 20, 127–130.

Van Doorslaer, K., Li, Z., Xirasagar, S., Maes, P., Kaminsky, D., Liou, D., et al., 2017. The Papillomavirus Episteme: a major update to the papillomavirus sequence database. Nucleic Acids Res. 45, D499–D506. https://doi.org/10.1093/nar/gkw879.

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., et al., 2017. Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8, 14049. https://doi.org/10.1038/ncomms14049.

Zhussupbekova, S., Sinha, R., Kuo, P., Lambert, P.F., Frazer, I.H., Tuong, Z.K., 2016. A mouse model of hyperproliferative human epithelium validated by keratin profiling shows an aberrant cytoskeletal response to injury. EBioMedicine 9, 314–323. https://doi.org/10.1016/j.ebiom.2016.06.011.