

Ermittlung von räumlichen Proteinexpressionsmustern mittels Bildverarbeitung

Protein-Profilung stratifizierter Epithelien mittels Bildverarbeitung fluoreszent gefärbter Gewebeschnitte

Thora Pommerencke¹, Pascal Tomakidi², Hartmut Dickhaus¹, Niels Grabe¹

¹Institut für medizinische Biometrie und Informatik, Universität Heidelberg

²Poliklinik für Kieferorthopädie der Universität Heidelberg

Email: thora.pommerencke@med.uni-heidelberg.de

Zusammenfassung. Wir präsentieren hier erste Ergebnisse einer neuartigen und vollautomatischen Methode zur Ermittlung von räumlichen Proteinexpressionsmustern in stratifizierten Epithelien. Das Verfahren basiert auf der Bildanalyse von immunhistologisch gefärbten Gewebeschnitten. Exemplarisch wird die Anwendbarkeit dieses Verfahrens anhand der Expression von fünf Strukturproteinen demonstriert.

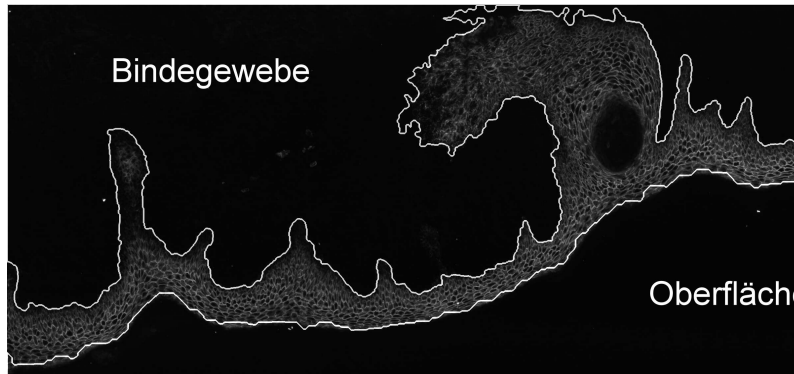
1 Einleitung

Äußere sowie innere Körperoberflächen werden von stratifiziertem, epithelalem Gewebe bedeckt. Dieses unterliegt einem komplex regulierten Gleichgewicht (Homöostase) aus Proliferation, Differenzierung und Apoptose bei einer über die Zeit konstanten Morphologie, wobei in der Literatur vier Schichten unterschieden werden: Stratum Basale, S. Spinosum, S. Granulosum und S. Corneum [1, 2]. Störungen der Homöostase führen zu starken pathologischen Veränderungen wie der Schuppenflechte und epithelialen Tumoren [3]. Grundlage für die Diagnose und Therapie dieser pathologischen Veränderungen ist ein tiefes Verständnis dieses komplexen biologischen Systems. In diesem Zusammenhang haben die systembiologische Analyse sowie Modellierung zunehmend an Bedeutung gewonnen [4, 5, 6, 7]. Da die epitheliale Homöostase wesentlich durch die gezielte Proteinexpression vieler Proteine gleichzeitig gesteuert wird, besteht zur dynamischen Modellierung des Epithels der Bedarf an multiparametrischen (d.h. mehrere Proteine werden simultan betrachtet) und quantitativen Daten zur Proteinexpression in Abhängigkeit von der Zelldifferenzierung.

2 Stand der Forschung und Fortschritt durch den Beitrag

Gängige Verfahren zur quantitativen Messung der Proteinexpressionsmuster in der Zelle/Gewebe (Protein-Profilung) auf Basis von Protein-Arrays und 2D-Gel-Elektrophorese basieren auf der Homogenisierung von Gewebe, wodurch jedoch

Abb. 1. Bild der Markerfärbung (Desmoplakin) mit detektierter Epithelkontur ohne Stratum Corneum



die Korrelation zwischen räumlicher Lage und Proteinexpression verloren geht [8, 9]. Neuere Verfahren wie z.B. das MALDI-Imaging bieten zwar eine Ortsauflösung, welche zur Zeit jedoch für die hier vorgestellte Anwendung noch zu gering ist [10]. Räumliche Beschreibungen der Proteinexpression entstammen in der Literatur daher hauptsächlich noch immer der mikroskopischen Analyse einzelner Gewebeschnitte durch einen Spezialisten. Die daraus resultierenden linguistischen und bildhaften Angaben lassen sich jedoch nur schwer in Parameter für Gewebesimulationen übersetzen [11]. Darüber hinaus sind die Daten unterschiedlicher Veröffentlichungen wegen mangelnder Standardisierung kaum vergleichbar. Aufgrund dieser unbefriedigenden Datenlage wurde ein neues Protein-Profilings-Verfahren entwickelt, welches die epitheliale Proteinexpression multiparametrisch und ortsauflösend durch automatische Bildanalyse gefärbter Gewebeschnitte quantifizieren kann.

3 Methoden

Experimentell wurden Schnitte cryokonservierter Gewebeproben von gesunden Patienten über eine fluoreszente Dreifachfärbung angefärbt. Dabei wird das interessierende Protein sowie das Kollagen-I als Bestandteil der extrazellulären Matrix im Bindegewebe über indirekte Immunfluoreszenz und die Zellkerne direkt chemisch mit DAPI (Diamidino-phenylindol) gefärbt. Da auf einem einzelnen Gewebeschnitt nur eine begrenzte Anzahl von Proteinen markiert werden können, wird die Analyse auf Serienschnitten durchgeführt, wodurch dieselbe Gewebeherkunft für jedes der Proteinexpressionsmuster gewährleistet ist.

Die Implementierung der Bildanalyse der digitalisierten Schnitte wurde mithilfe von Matlab 7.1 einschließlich der Image Processing Toolbox realisiert. In einem ersten Schritt wird das Epithel vollautomatisch segmentiert, wobei die Gewebekonturen über ein Phasenkontrastbild ermittelt werden. Die Abgrenzung des Epithels vom Bindegewebe erfolgt im wesentlichen über eine gezielte Canny-Kantendetektion im Bild der Zellkernfärbung sowie der Auswahl von epithelialen

Gewebebereichen anhand der Bindegewebsfärbung nahe den Kanten. Von der ermittelten Epithelmaske werden Bereiche starker Bindegewebsfärbung durch logisches NOT ausgeschlossen. Das segmentierte Epithel wird anschließend künstlich um etwa eine halbe Zellbreite in Richtung Bindegewebe dilatiert, um auch den zellfreien Raum zwischen Epithel und Bindegewebe, die Basalmembran, mit in die Analysen einschließen zu können. Das Stratum Corneum lässt sich über die Abwesenheit von Zellkernen ermitteln und gegebenenfalls ausschließen.

Die funktionale Abhängigkeit der Proteinexpression von der Differenzierung wird von uns über zwei verschiedene Maße beschrieben. Zum einen über die normierte Distanz d zum Bindegewebe (mit $d = 0\%$ direkt am Bindegewebe und $d = 100\%$ an der Oberfläche), da sich differenzierende Zellen in Richtung Oberfläche bewegen. Zum anderen über die mittlere Intensität der Zellkernfärbung, welche die mittlere Zellkerndichte widerspiegelt. Diese nimmt durch die Vergrößerung der Zellen sowie die Auflösung des Zellkernes im Zuge der terminalen Differenzierung ab. Für das zellkernbasierte Maß wurde ein Glättungsverfahren entwickelt, bei dem das Bild der Zellkernfärbung mit einer annähernd rechteckigen Filtermatrix parallel zur Basalschicht gefaltet wird, dessen Breite und Höhe mit steigender Distanz zum Bindegewebe zunehmen (siehe Gleichung 1).

$$G(p) = \frac{\sum_{i=1}^{|N(p)|} E(N(p)_i)}{|N(p)|}$$

$$N(p) = \left\{ x \in M \mid M(x) \neq 0 \wedge d(x, p) \leq \frac{l_p}{2} \wedge |d(x, b(x)) - d(p, b(p))| \leq \frac{l_p}{4} \right\} \quad (1)$$

$$l_p = f \cdot d(p, b(p)) + l_{min}$$

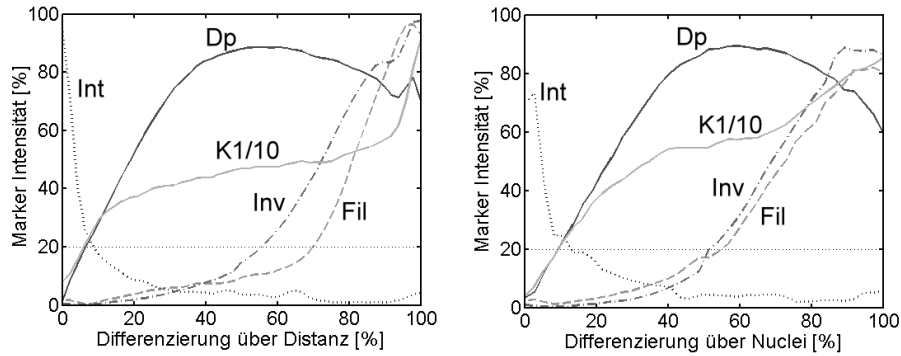
E: Bild der Zellkernfärbung	G: geglättetes Bild
M: Epithelmaske	p: aktuell betrachtetes Pixel
f: Wachstumsfaktor	l_{min} : Mindestkantenlänge der Filtermatrix
I(x): Wert von Pixel x in Bild I	l_p : aktuelle Kantenlänge der Filtermatrix
N(p): Menge der Nachbarschaftspixel von p	
d: Distanzfunktion, $d(x, p)$ liefert euklidische Distanz zwischen Pixel x und p	
b: $b(p)$ liefert nächstes Bindegewebspixel zu p	

Auf Basis dieser Differenzierungsmaße kann zu jeder „Differenzierungsstufe“ die jeweilige mittlere Färbungsintensität des Markers ermittelt werden. Die Auftragung der normierten Markerintensität gegen das Differenzierungsmaß ergibt für jeden Marker ein charakteristisches Expressionsprofil.

4 Ergebnisse

Mit dem entwickelten Verfahren wurde hier beispielhaft die kombinatorische Expression von den fünf Strukturproteinen Desmoplakin (Dp), Filaggrin (Fil), Integrin- $\alpha_6\beta_4$ (Int), Involucrin (Inv) und Keratin 1/10 (K1/10) untersucht, die

Abb. 2. Generierte Multi-Marker-Profile der humanen Epidermis ohne Stratum Corneum über den distanzbasierten Ansatz (links) und über den zellkernbasierten Ansatz (rechts)



charakteristisch für die üblicherweise beschriebenen „Differenzierungsstufen“ sind. Für die Analyse wurde eine epidermale Gewebeprobe (humane Haut) gewählt. In dem dargestellten Multi-Marker-Profil handelt es sich je Marker um ein gemittelttes Expressionsprofil aus mehreren Bilddatensätzen derselben Gewebeprobe.

5 Diskussion

Das realisierte Profiling-Verfahren ist in der Lage, erstmals automatisiert eine grobe quantitative und räumliche Beschreibung der Proteinexpression im Epithel zu liefern. Die ermittelten epidermalen Expressionsprofile stimmen mit den in der Literatur qualitativ beschriebenen überein. Für die Generierung der Profile sind die zugrunde liegenden „Differenzierungsstufen“ beliebig groß oder klein wählbar. Dadurch kann eine sehr hohe räumliche Auflösung der Profile erzielt werden, welche eine kontinuierlichen Messung ohne a priori interpretative Unterteilung in die herkömmlichen vier Schichten.

Die beiden hier gewählten Differenzierungsmaße stellen eine erste Realisierung dar. Besonders in Bezug auf die Schärfe der Profile sind jedoch Verbesserungen denkbar. Inwieweit die Beschreibung des Differenzierungsgrades über die mittlere Anzahl an Zellkernen pro Fläche einen Vorteil gegenüber der Korrelation von Differenzierung und Distanz zum Bindegewebe bringt, muss durch die Analyse weiterer Datensätze überprüft werden. In den bisherigen Ergebnissen zeigt der zellkernbasierte Ansatz eine geringere statistische Sicherheit und ein schlechteres Auflösungsvermögen, doch ist bei komplexeren räumlichen Epithelstrukturen mit einer gestörten Differenzierungsschichtung ein robusteres Verhalten gegenüber dem distanzbasierten Ansatz zu erwarten.

Die vorgestellte Methodik zur Generierung von Expressionsprofilen bildet eine erste Grundlage für mathematische Beschreibungen der räumlichen Proteinexpression in stratifizierten Epithelien. Zum einen könnten die gemessenen funk-

tionalen Abhängigkeiten in ein systembiologisches Modell der epithelialen Homöostase integriert werden. Zum anderen würden die ermittelten Beschreibungen der Expression einen statistischen Vergleich von Gewebeproben unterschiedlicher Gewebelokalisationen bzw. von pathologischem und Normalgewebe ermöglichen.

6 Danksagung

Die Bildakquise wurde am Nikon Imaging Center in Heidelberg durchgeführt.

Literaturverzeichnis

1. Candi E, Schmidt R, Melino G. The cornified envelope: A model of cell death in the skin. *Nature* 2005;6(4):328–338.
2. Fuchs E, Byrne C. The epidermis: Rising to the surface. *Curr Opin Cell Biol* 1994;4(5):725–736.
3. Carlos L, Junqueira U, Carneiro J. *Histologie*. Springer, Berlin; 2005.
4. Morel D, Marcelpoil R, Brugal G. A proliferation control network model: The simulation of two-dimensional epithelial homeostasis. *Acta Biotheoretica* 2001;49(4):219–234.
5. Rashbass J, Stekel D, Williams E. The use of a computer model to simulate epithelial pathologies. *J Pathology* 1996;179(3):333–9.
6. Walker D, Southgate J, et al GHill. Agent-based computational modelling of epithelial cell monolayers. *IEEE/ACM Trans Nanobioscience* 2004;3(3):153–163.
7. Grabe N, Neuber K. A multicellular systems biology model predicts epidermal morphology, kinetics and Ca²⁺ flow. *Bioinformatics* 2005;21(17):3541–7.
8. Farrell P. High resolution two-dimensional electrophoresis of proteins. *J Biological Chem* 1975;250(10):4007–21.
9. Fung E, Thulasiraman V, Weinberger S. Protein biochips for differential profiling. *Curr Opin Biotech* 2001;12(1):65–9.
10. Chaurand P, Sanders M, et al RJensen. Proteomics in diagnostic pathology: Profiling and imaging proteins directly in tissue sections. *Am J Pathol* 2004;165(4):1057–68.
11. Murphy R. Location proteomics: A systems approach to subcellular location. *Biochem Soc* 2005;33(3):535–8.