SCIENTIFIC DATA

OPEN

Data Descriptor: RNA-seq of macrophages of amoeboid or mesenchymal migratory phenotype due to specific structure of environment

Vladimír Čermák[1,2], Aneta Gandalovičová[1,2], Ladislav Merta[1,2], Jitka Fučíková[3,4], Radek Špíšek[3,4], Daniel Rösel[1,2] & Jan Brábek[1,2]

M2-polarized macrophages have been shown to adapt their 3D migration mode to physical properties of surrounding extracellular matrix. They migrate in the integrin-mediated adhesion and proteolytic activity-dependent "mesenchymal" mode in stiff matrices and in the integrin and protease-independent "amoeboid" mode in low density, porous environments. To find out what impact the switching between the migration modes has on expression of both protein-coding and non-coding genes we employed RNA sequencing of total RNA depleted of ribosomal RNA isolated from macrophages migrating in either mode in 3D collagens. Differentially expressed genes from both categories have been detected and the changes in expression of selected genes were further validated with RT-qPCR. The acquired data will facilitate better understanding of how mechanical properties of tissue microenvironment reflect in macrophage immune function and how the transitions between mesenchymal and amoeboid migratory modes are regulated at the gene expression level.

Design Type(s)	factorial design • transcription profiling design
Measurement Type(s)	transcription profiling assay
Technology Type(s)	RNA sequencing
Factor Type(s)	Blood Donor • collagen-containing extracellular matrix
Sample Characteristic(s)	Homo sapiens • macrophage

[1]Department of Cell Biology, Charles University, Viničná 7, Prague, Czech Republic. [2]Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University (BIOCEV), Průmyslová 595, 25242, Vestec u Prahy, Czech Republic. [3]Department of Immunology, Charles University, 2nd Faculty of Medicine and University hospital Motol, Prague, Czech Republic. [4]Sotio, Prague, Czech Republic. Correspondence and requests for materials should be addressed to J.B. (email: jan.brabek@natur.cuni.cz).

Background & Summary

Cell invasion is required for many physiological processes such as development, immune response or wound healing, and, in effect, many cell types possess the capability to invade[1]. Immune cells have developed very effective invasion strategies that enable them to perform immune surveillance even deep in tissues. However, cell invasion also facilitates dangerous pathological states such as cancer metastasis, and cancer cells are often considered to be masters of invasion. A key feature underlying effective invasion, shared by both immune cells and cancer cells, is the ability to adapt to different conditions by adjusting the mode of invasion, often referred to as invasion plasticity[2–5].

The various invasion modes cells can adopt include two distinct individual invasion modes – the protease-dependent mesenchymal mode and protease-independent amoeboid mode. Generally, mesenchymal migration relies on the formation of adhesions to the extracellular matrix (ECM) and proteolytic degradation of adjacent fibres[6,7]. On the contrary, amoeboid cells do not rely on fibre degradation and instead dynamically deform their cell body (including the nucleus), to squeeze through pores in the ECM[8,9]. These invasion modes are interchangeable and cells can undergo the mesenchymal-amoeboid (MAT) or amoeboid-mesenchymal transitions (AMT) in response to intracellular and/or extracellular cues[6,10–14]. While processes of MAT and AMT in cancer cells have been extensively studied during the past years, little information is available on the molecular mechanisms behind the invasion plasticity of immune cells[3,15].

An elaborate study by Cougole et al. tested the invasion plasticity of various immune cells in response to different ECM conditions and showed that these cells differ in their capability to adjust to soft or stiff matrices. They conclude that all leukocytes use the amoeboid mode, but only M2 macrophages possessed the ability to adjust their morphology according to the extracellular matrix conditions and switch between the amoeboid and mesenchymal phenotype[15].

This phenomenon was described in detail by Van Goethem et al., which showed that M2 macrophages modulate their invasion mode according to the ECM architecture[3]. In Matrigel and gelled collagen, both dense environments with small pores, M2 macrophages utilize the mesenchymal mode and form podosomes - membrane protrusions with proteolytic activity. Interestingly, this invasion mode can be inhibited by a protease inhibitor mix, but not by a selective inhibitor of matrix metalloproteinases alone, indicating other proteases to be involved. Moreover, protease inhibition does not result in MAT, as described in cancer cells. In fibrillar collagen, which permits cells to take advantage of larger pre-existing pores, macrophages are amoeboid and do not form podosomes. Amoeboid macrophage invasion is unaffected by protease inhibitors but is decreased in the presence of ROCK inhibitors, which is in line with observations of amoeboid cancer cells.

To elucidate signalling driving this invasion adaptability of macrophages and potentially reveal candidate marker genes of the amoeboid and mesenchymal phenotype, we performed RNA sequencing of three independent primary samples of human M2-polarized macrophages embedded in both low and high-density collagen (Fig. 1), as described by Van Goethem et al.[3] In agreement with their results, if we embedded M2 macrophages in 1.7 mg/ml collagen, they adopted the amoeboid phenotype in over 86% (SE 4.36%). In dense collagen (4.8 mg/ml), the majority (83%, SE 2.65%) of macrophages was mesenchymal (Fig. 2). To verify the migratory potential of both amoeboid and mesenchymal macrophage invasion modes, we performed time-lapse live cell imaging. Supplementary videos clearly demonstrate amoeboid locomotion in fibrillar collagen (Macrophage phenotype in fibrillar colagen, Data Citation 1) and document mesenchymal invasion of M2 macrophages in high-density environment (Macrophage phenotype in dense collagen, Data Citation 1).

A substantial advantage of primary M2 macrophages as a model system is their inherent invasion plasticity that enables to study mechanisms of MAT and AMT without any genetic manipulation or chemical treatment. The presented data will help to better comprehend the in vivo plasticity of cellular invasion caused by ECM diversity. In addition, analysis of the data revealed that the immune function of macrophages is also affected by mechanical properties of the ECM, hence reuse of the data by the immunological community could be beneficial.

Methods

Differentiation and M2 polarization of primary human macrophages

Primary human macrophages derived from CD14[+] subpopulation of peripheral blood mononuclear cells from healthy donors aged between 25–40 years were differentiated and M2-polarized with M-CSF (100 ng/ml) in RPMI-1640 medium supplemented with 10% fetal bovine serum for 5–6 days before the experiments. Informed consent was obtained from all subjects. The experiments were approved by the Ethical Committee of the Second Faculty of Medicine at Charles University in Prague.

Three-dimensional cell culture and RNA isolation

Five million M2 macrophages were cultured in 500 µl of three-dimensional gel made of bovine skin collagen (Nutragen, Advanced BioMatrix, 1.7 or 4.8 mg/ml) in RPMI-1640 with 1% fetal bovine serum and 100 ng/ml M-CSF (Peprotech) for 48 hours in a 24-well plate. Gels from two wells were transferred into one 2 ml tube and homogenized with Tissue Tearor (BioSpec Products) in 600 µl of RNA extraction solution (60% v/v water-saturated phenol, 3.25 M guanidine thiocyanate, 400 mM sodium acetate buffer pH 4.0, 0.4% w/v N-lauroylsarcosine, 160 mM 2-mercaptoethanol) plus 100 µl of 6.1 M sodium chloride.
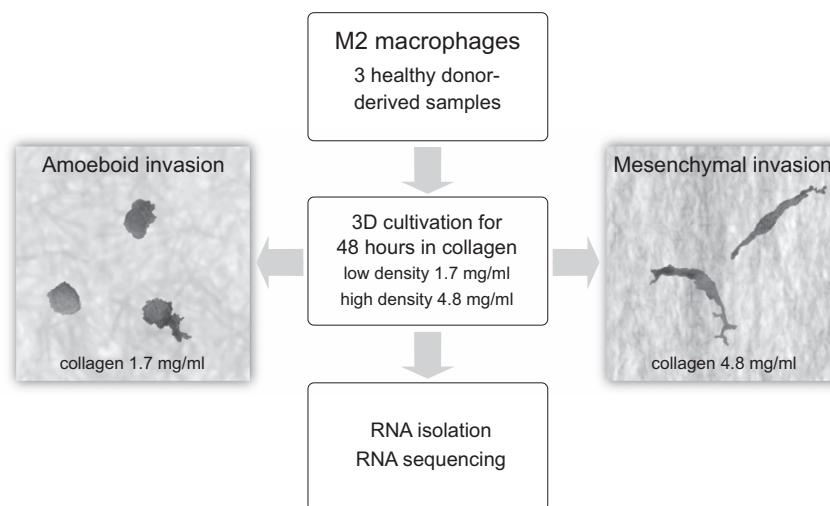
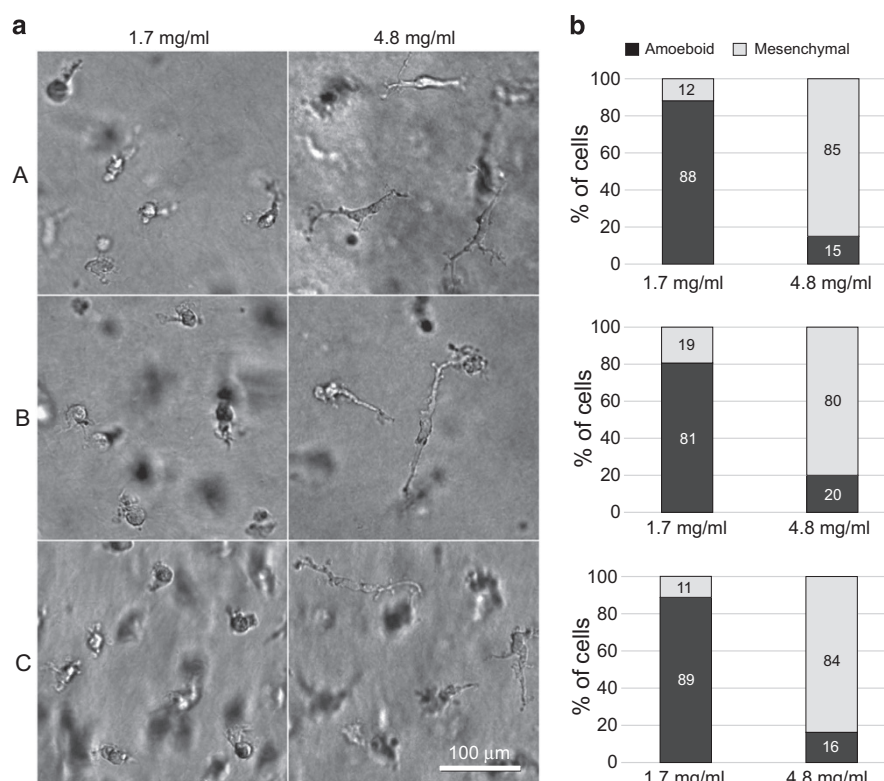**Figure 1.** Schematic overview and experimental design of the study.

Figure 2. **Morphology analysis of M2 macrophage samples embedded in low- and high-density collagen.** (**a**) Representative images of the three analysed samples of M2 macrophages (A–C) embedded in 1.7 mg/ml or 4.8 mg/ml collagen. Images were acquired using Nikon-Eclipse TE2000-S, 10x objective with Hoffman modulation contrast. (**b**) Quantification of cell phenotype of the three analysed samples of M2 macrophages (a–c) embedded in 1.7 mg/ml or 4.8 mg/ml collagen.

200 µl of chlorophorm was added to approx. 1 ml of the lysate and the mixture was vortexed vigorously for 10 seconds. After 30-minute centrifugation at 18,000 g at 4 °C, the polar upper phase was transferred into a new tube, volume adjusted to 800 µl with RNase-free water, RNA was precipitated with 600 µl of isopropanol at −20 °C overnight and recovered by centrifugation at 18,000 g for 30 minutes at 4 °C. The RNA pellet was washed three times with 600 µl of 75% ethanol and air-dried. Next, the RNA was treated with DNase I to remove possible genomic DNA contamination. To this end, the pellet was

Subjects	Gender	Condition	Protocols	ENA	BioSD
Donor A	male	1.7 mg/ml collagen	P-MTAB-73510 to P-MTAB-73516	ERS2359742	SAMEA1054135
Donor A	male	4.8 mg/ml collagen	P-MTAB-73510 to P-MTAB-73516	ERS2359739	SAMEA1054132
Donor B	female	1.7 mg/ml collagen	P-MTAB-73510 to P-MTAB-73516	ERS2359743	SAMEA1054136
Donor B	female	4.8 mg/ml collagen	P-MTAB-73510 to P-MTAB-73516	ERS2359740	SAMEA1054133
Donor C	male	1.7 mg/ml collagen	P-MTAB-73510 to P-MTAB-73516	ERS2359744	SAMEA1054137
Donor C	male	4.8 mg/ml collagen	P-MTAB-73510 to P-MTAB-73516	ERS2359741	SAMEA1054134

Table 1. Summary of sequencing data.

directly dissolved in 100 µl of a solution containing 4 units of DNase I (Thermo Fisher Scientific) and manufacturer-provided reaction buffer, and incubated at 37 °C for 30 minutes. After that, the RNA was re-purified with RNeasy Protect Mini Kit (Qiagen) according to the manufacturer's instructions and eluted in 50 µl of RNase-free water (Invitrogen).

Analysis of cell morphology in 3D collagen

Cell morphology was analysed with ImageJ software by measuring longest axis to shortest axis ratios. Cells with the ratio equal to 2.5 or larger were considered mesenchymal, as determined earlier by Van Goethem et al.[3]. For each condition a minimum of 100 cells was analysed.

RNA sequencing and data processing

Stranded, Illumina HiSeq-compatible library was constructed with ScriptSeq Complete (Human/Mouse/Rat) library preparation kit (Epicentre) according to the manufacturer's instructions. An equimolar pool of 6 sample libraries (three donor-matched pairs) was sequenced on one whole lane of Illumina HiSeq 2500 sequencer in high output, paired (2×125 cycles) mode by Beckman Coulter Genomics.

Raw reads were trimmed of adapter sequences with Cutadapt[16] (version 1.15) and mapped to human genome version GRCh38.91 with the STAR short read aligner[17] version 2.5.4b with default settings and output extended with read counts per gene.

Adapter-trimmed reads were deposited in the ArrayExpress database (Data Citation 2). The samples and related files are summarized in Table 1.

Differential Gene Expression Analysis

To find differentially expressed genes R package DESeq2[18] (version 1.18.1) was used with the raw read count output of the STAR aligner as the input. Default workflow with a design formula reflecting paired nature of the samples was applied. Differentially expressed genes with FDR < 0.1 (874 upregulated and 1,676 downregulated genes) are available in the file Differentially_expressed_genes.xlsx (Data Citation 1) and their distribution by gene expression level is depicted as MA plot in Fig. 3d. Principle component analysis of the gene expression profiles showed dominant clustering of the samples by individual donors (Fig. 3e). This obviously reflects the well-known high variation of primary cell samples that can much exceed the variation brought about by experimental treatment as in this case. The use of two-factor design (cell donor and collagen concentration) in the statistical analysis is thus crucial.

Analysis of ribosomal cDNA content and strandedness of libraries

Adapter-trimmed reads were mapped to human ribosomal RNA coding sequences (NR_023379, NR_003285, NR_003286, NR_003287, NR_137294, and NR_137295) with BBMap aligner[19] in local alignment mode. Ribosomal cDNA content in the libraries was expressed as percentage of reads mapped to nuclear or mitochondrial rRNA transcripts (the larger number from the paired mates). Strandedness of libraries was assessed from ratios of read counts mapped (by STAR) to annotated genomic loci in sense and antisense orientation. To eliminate the noise associated with low-expressed genes we limited the calculations to genes with lognormal mean or higher expression (approx. 6,000 genes). We also excluded genes with more antisense transcripts than those in sense orientation. The mean sense/antisense ratio was calculated as a geometric mean of all included gene ratios across all samples.

Reverse Transcription – Quantitative polymerase chain reaction (RT-qPCR)

All the RT-qPCR experiments were performed according to MIQE guidelines[20]. Briefly, total RNA was extracted from cells embedded in 3D collagen matrix as described above. RNA reverse transcription was performed using M-MuLV Reverse Transcriptase (NEB) with 1 µM oligo-dT and 0.2 µM MALAT1-specific primer and 700 ng of total RNA. RT-qPCR was performed using SYBR green mix (1x Standard Taq buffer (NEB) with 3.0 mM MgCl$_2$, 2 U/100 µl Hot Start Taq DNA Polymerase (NEB), 10% glycerol (Carl Roth), 6% DMSO (Sigma), 200 µM dNTPs (Thermo Fischer Scientific), 0.05% Triton X-100 (Sigma), 1x SYBR® Green I Nucleic Acid Stain (Lonza), and Ultrapure Water (Invitrogen)) and CFX384 Real Time PCR Instrument (Bio-Rad). Annealing and elongation temperatures were 63 °C and 68 °C, respectively. The RT-qPCR reaction volume was 10 µl with 0.5 µl of cDNA and 200 nM primers (for
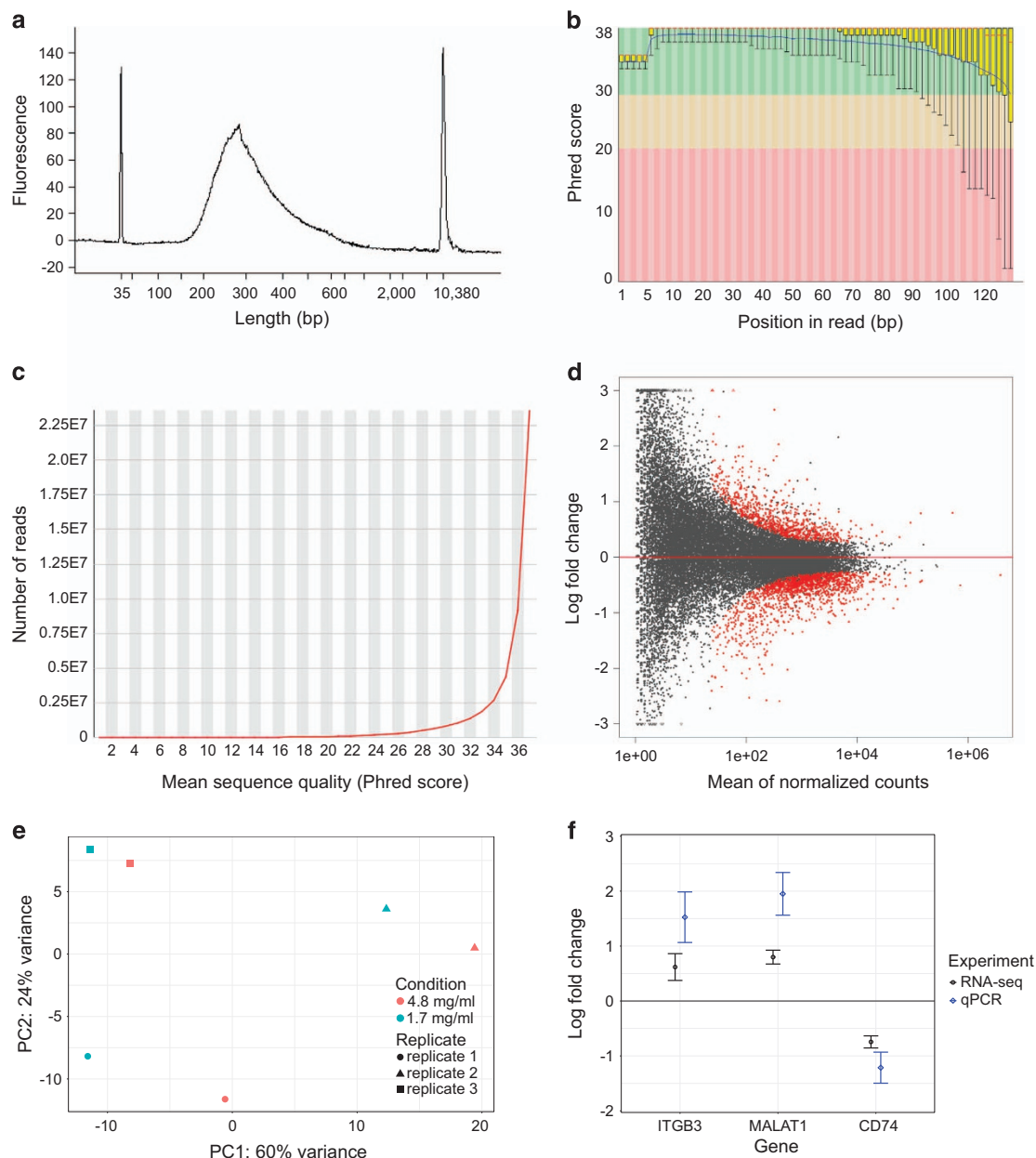
Figure 3. RNA-seq results validation. (**a**) Sequencing library fragment size distribution of sample 1. (**b**) Per base sequence quality expressed as Phred score by position, sample 1, first reads. (**c**) Quality score distribution over all reads of sample 1. (**d**) MA plot of log2 fold change values against normalized counts for each gene in the analysis. Red points mark genes with FDR < 0.1. (**e**) Principal component analysis of gene expression profiles. (**f**) Comparison of log2 fold change values detected by RNA-seq and RT-qPCR, respectively for the indicated genes.

primer details see Table 2). Samples were run as technical quadruplicates using FrameStar® 480/384 with RT-qPCR Adhesive seal (4titude) 384-well plates. Cq and relative expression values were calculated by setting single threshold value for each target and further analysed by qBase + software (Biogazelle)[21] using EIF4H, HNRNPL and PPIA as reference gene index (based on geNORM analysis[22]). Amplification efficiencies were determined by the standard curve analysis. Log2 fold change values were calculated from relative expression values.

Data Records

Time-lapse movies documenting migratory phenotypes of the cells and a table listing all the differentially expressed genes detected in the study were deposited in the Figshare repository (Data Citation 1).

RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (Data Citation 2).

Transcript	Acc. Number	Primer	Sequence	Product (bp)	Eff. (%)
MALAT1	NR_002819	MALAT1_5s	GAATTGCGTCATTTAAAGCCTA	85	92.2
		MALAT1_5a	GTTTCATCCTACCACTCCCAAT		
ITGB3	NM_002211	ITGB3f	CACCATCCACGACCGAAAAG	82	99.5
		ITGB3r	CAGTGGGTTGTTGGCTGTGT		
CD74	NM_001025159	CD74f	GGAAGGTCTTTGAGAGCTGGAT	114	111.0
		CD74r	CTTCCTGGCACTTGGTCAGTA		
HNRNPL	NM_001533	HNRNPLs	CTGGAGGTGACCGAGGAGAA	96	99.3
		HNRNPLa	GCGCTCACTTTTGCCTGAGAA		
EIF4H	NM_022170	EIF4Hs	CTTCGACACCTACGACGATCG	95	97.7
		EIF4Ha	CCTTCTGGCTACGGGAACCAT		
PPIA	NM_021130	PPIAf	GCCGAGGAAAACCGTGTACTA	106	94.1
		PPIAr	CTGCAAACAGCTCAAAGGAGAC		

Table 2. **Primers used for RT-qPCR.**

Sample		Nuclear-encoded rRNA	Mitochondrial rRNA
Donor A	1.7 mg/ml collagen	0.9%	9.5%
	4.8 mg/ml collagen	0.7%	11.8%
Donor B	1.7 mg/ml collagen	0.3%	9.8%
	4.8 mg/ml collagen	0.4%	16.0%
Donor C	1.7 mg/ml collagen	1.1%	11.0%
	4.8 mg/ml collagen	0.6%	13.5%

Table 3. **Percentage of rRNA matching reads in RNA-seq data.**

Technical Validation

Viability and morphological phenotype of the cells was verified by wide field microscopy (Fig. 2). The integrity of isolated total RNA was confirmed with agarose gel electrophoresis of denatured RNA samples. The quality and size distribution of sequencing libraries was analysed with Agilent BioAnalyzer 2100 (Fig. 3a). The profiles, according to expectation, showed peak position around 280 bp and a longer right tail. The output statistics of STAR aligner showed an average paired read length of 230 bases. Raw reads trimmed of adapter sequences, approx. 2 × 40 million reads per sample, were quality-checked with FastQC software[23]. FastQC's plots of Phred scores by position (Fig. 3b) showed typical profiles with decreasing quality towards the ends of reads. The overall quality was analysed with BBTools' Reformat utility[19] detecting approx. 91% bases having Phred score 30 or better. The distribution of quality scores is presented in Fig. 3c. Figure 3 displays quality assessment results for sample 1. Complete results for all six samples are provided in Supplementary Figure 1. FastQC analysis also detected a significant presence of duplicate reads, which was expected as the library construction involved a PCR amplification step. The STAR aligner uniquely mapped approx. 90% of the fragments to human genome, approx. 6% fragments were multi-mapped, and approx. 4% fragments were excluded as too short. We further analysed residual rRNA content after ribodepletion expressed as proportion of reads mapped to rRNA sequences. Only 0.3–1.1% reads could be mapped to nuclear encoded rRNA. However, 9.5–16% of all reads were mapped to mitochondrial 12S and 16S rRNA sequences with higher mtrRNA content in all high-concentration collagen samples (Table 3). The Ribo-Zero component of the library construction kit used in this work does not deplete total RNA of mitochondrial rRNA, so this is not unexpected.

The libraries were prepared as strand-specific. To analyse the level of strandedness in the sequencing data we compared proportions of sense and antisense reads mapped to known genes of intermediate to high expression. The average sense/antisense ratio was 18.4 ± 1.87 (geometric mean and standard deviation). This might appear slightly less than expected, however a few factors lowering apparent strandedness should be taken into account. The ribodepleted total RNA is much more complex than mRNA isolated with oligo-dT probes. It contains unprocessed primary transcripts that may partly span neighbouring genes in opposite orientation. It also contains large proportion of non-coding sequences

that sometimes arise from antisense transcription of coding genes. Another source of partial loss of strandedness is the DNA-dependent DNA polymerase activity of reverse transcriptase.

To validate the results of differential gene expression analysis, RT-qPCR was performed with the RNA samples used for sequencing to compare changes in expression of three selected genes – ITGB3, MALAT1 and CD74 as detected by RNA-seq and RT-qPCR. These genes were chosen in order to represent different functional categories (cell adhesion, antigen presentation) and also both protein-coding and non-coding (MALAT1) genes. The results showed essentially the same expression pattern with RT-qPCR detecting greater fold-changes than RNA-seq (Fig. 3f).

References

1. Binamé, F., Pawlak, G., Roux, P. & Hibner, U. What makes cells move: requirements and obstacles for spontaneous cell motility. *Mol. Biosyst.* **6,** 648–661 (2010).
2. Friedl, P. & Alexander, S. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell* **147,** 992–1009 (2011).
3. Van Goethem, E., Poincloux, R., Gauffre, F., Maridonneau-Parini, I. & Le Cabec, V. Matrix architecture dictates three-dimensional migration modes of human macrophages: differential involvement of proteases and podosome-like structures. *J. Immunol.* **184,** 1049–1061 (2010).
4. Taddei, M. L., Giannoni, E., Comito, G. & Chiarugi, P. Microenvironment and tumor cell plasticity: an easy way out. *Cancer Lett.* **341,** 80–96 (2013).
5. Charras, G. & Sahai, E. Physical influences of the extracellular environment on cell migration. *Nat. Rev. Mol. Cell Biol.* **15,** 813–824 (2014).
6. Wolf, K. *et al.* Compensation mechanism in tumor cell migration: mesenchymal-amoeboid transition after blocking of pericellular proteolysis. *J. Cell Biol.* **160,** 267–277 (2003).
7. Nagano, M., Hoshino, D., Koshikawa, N., Akizawa, T. & Seiki, M. Turnover of focal adhesions and cancer cell migration. *Int. J. Cell Biol.* **2012,** 310616 (2012).
8. Lämmermann, T. & Sixt, M. Mechanical modes of 'amoeboid' cell migration. *Curr. Opin. Cell Biol.* **21,** 636–644 (2009).
9. Charras, G. & Paluch, E. Blebs lead the way: how to migrate without lamellipodia. *Nat. Rev. Mol. Cell Biol.* **9,** 730–736 (2008).
10. Friedl, P. & Wolf, K. Plasticity of cell migration: a multiscale tuning model. *J. Cell Biol.* **188,** 11–19 (2010).
11. Liu, Y. J. *et al.* Confinement and low adhesion induce fast amoeboid migration of slow mesenchymal cells. *Cell* **160,** 659–672 (2015).
12. Paňková, K., Rösel, D., Novotný, M. & Brábek, J. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cell. Mol. Life Sci.* **67,** 63–71 (2010).
13. Paňková, D. *et al.* NG2-mediated Rho activation promotes amoeboid invasiveness of cancer cells. *Eur. J. Cell Biol.* **91,** 969–977 (2012).
14. Ladhani, O., Sanchez-Martinez, C., Orgaz, J. L., Jimenez, B. & Volpert, O. V. Pigment epithelium-derived factor blocks tumor extravasation by suppressing amoeboid morphology and mesenchymal proteolysis. *Neoplasia* **13,** 633–642 (2011).
15. Cougoule, C. *et al.* Blood leukocytes and macrophages of various phenotypes have distinct abilities to form podosomes and to migrate in 3D environments. *Eur. J. Cell Biol.* **91,** 938–949 (2012).
16. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).
17. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
18. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).
19. Bushnell, B. BBTools software package. BBTools - DOE Joint Genome Institute https://jgi.doe.gov/data-and-tools/bbtools (2014).
20. Bustin, S. A. *et al.* The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **55,** 611–622 (2009).
21. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **8,** R19 (2007).
22. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3,** research0034.1–0034.11 (2002).
23. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinformatics* https://www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).

Data Citations

1. Čermák, V. *et al. figshare* https://doi.org/10.6084/m9.figshare.c.4140770 (2018).
2. *ArrayExpress* E-MTAB-6643 (2018).

Author Contributions

V.Č. prepared sequencing libraries and analysed the data. A.G. performed 3D cell culture experiments. L. M. performed RT-qPCR experiments. J.F. prepared human donor-derived macrophages. R.Š. conceived the study and supervised experiments. D.R. conceived the study and supervised experiments. J.B. conceived the study and supervised experiments. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/sdata

Competing interests: The authors declare no competing interests.

How to cite this article: Čermák, V. *et al.* RNA-seq of macrophages of amoeboid or mesenchymal migratory phenotype due to specific structure of environment. *Sci. Data*. 5:180198 doi: 10.1038/sdata.2018.198 (2018).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.