

RESEARCH

Open Access



Cumulative prognostic power of laminin genes in colorectal cancer

Vladimir V. Galatenko^{1,2,3}, Diana V. Maltseva², Alexey V. Galatenko¹, Sergey Rodin^{2,4*†} and Alexander G. Tonevitsky^{2,5*†}

From Belyaev Conference
Novosibirsk, Russia. 07-10 August 2017

Abstract

Background: Laminins are a major family of extracellular matrix proteins and the main component of basement membranes. Laminins are involved in many if not all stages of cancer progression, and expression of laminin genes has prognostic value in various types of cancer, including colorectal. Only single laminin genes or components of a single laminin trimer with significant differential expression have been regarded as potential biomarkers to date.

Results: Here we compared prognostic power of classifiers constructed from sets of laminin genes with that of any single laminin gene. The analysis showed that cumulative prognostic power of sets of laminin genes was higher and was achieved already with pairs and triples of the genes. Interestingly, components of the pairs and the triples did not belong to any known laminin trimer, but, taken together with the gene weights, suggested higher *LAMA4/LAMA5* expression ratio in patients with poor prognosis.

Conclusions: Analysis of the laminin expression profile rather than expression of the single genes or components of laminin trimers is useful for colorectal cancer prognosis in patients. High *LAMA4/LAMA5* ratio is associated with increased permeability of basement membranes suggesting that basement membranes produced by colorectal tumors might be an important hindrance to their own dissemination in patients.

Keywords: Laminins, Cancer prognosis, Colorectal cancer

Background

Laminins are secreted, multidomain heterotrimeric proteins that consist of one α , one β , and one γ chains [1]. Five laminin α , four β , and three γ chains, which form not less than 16 different trimers, have been discovered in humans to date [2]. Unlike the majority of extracellular matrix proteins, laminins exhibit a certain degree of tissue and temporal specific expression [3]. As the main component of basement membranes (BMs), laminins affect various properties of BMs, for instance their permeability to cells.

Laminins are significantly involved in survival and proliferation of tumors and cancer cells, angiogenesis, migration and breaching of BMs by cancer cells, development of pre-metastatic niches at distant organs and many other

stages of cancer progression (recently reviewed in [4]). A plethora of studies has demonstrated prognostic value of laminins in a range of cancers, including colorectal [4, 5]. However, these studies were focused only on expression of single laminin genes [6–8] or, rarely, components of a single laminin trimer [9]. Meanwhile, it has been shown that prognostic power of a set of genes can be higher than a sum of prognostic powers of the individual genes from the set [10]. Here we sought to identify sets of laminins with higher cumulative prognostic power than single laminin chains in colorectal cancer.

Methods

Estimation of prognostic power of laminin genes

Estimation of cumulative prognostic power of laminin genes was performed according to a previously described method [10]. In short, the estimation included construction classifiers based on expression levels of laminin genes

* Correspondence: sergey.rodin@ki.se; tonevitsky@mail.ru

†Equal contributors

²SRC Bioclinicum, Ugreshskaya str 2/85, 115088 Moscow, Russia
Full list of author information is available at the end of the article

(all laminin genes or subsets of laminin genes), filtering out low-quality classifiers, and evaluation of quality of these classifiers. Classifier construction was performed using soft-margin support vector machine (SVM) [11] with linear kernel applied to a training dataset (see Microarray datasets). At the filtration stage quality of the classifiers was evaluated independently on two filtering datasets (see Microarray datasets). The area under receiver operating characteristic (ROC) curve (AUC – area under ROC curve) was used as the main characteristic of the classifiers quality. A classifier passed filtration stage if AUC values were not less than 0.6 for the training dataset and not less than 0.55 for each filtration dataset. Final quality evaluation was performed on a testing dataset (see Microarray datasets) that was a totally independent dataset and contained no data intersections with the training or the filtering datasets.

Microarray datasets

The utilized meta-set included GSE39582 [12] dataset that was used for training; GSE37892 [13] and GSE17536 [14, 15] datasets used for filtration; and GSE14333 [16] dataset used for validation of the results (testing dataset). The analyzed datasets (which were all based on Affymetrix Human Genome U133 Plus 2.0 Arrays) were jointly preprocessed using Robust Multi-array Average (RMA) method [17]. The patients with unknown recurrence status or with recurrence during the first month of follow-up were excluded from the analysis. Information on the number of patients in each dataset is presented in Table 1.

Only patients with a recurrence during the first 3 years and patients without a diagnosed recurrence and at least 4 years of follow-up were used for training and filtration as well as for construction of the ROC-curves and calculation of AUC for the testing dataset. Kaplan-Meier curves were plotted based on the complete testing dataset, including information on patients without a recurrence but time of follow-up less than 4 years and patients with recurrence but at least 3 years recurrence-free survival.

If the array contained several probes for a laminin chain, only one probe (with the highest intensity of the signal) was taken for the analysis. Laminin genes *LAMB4* and *LAMC3* were excluded from the analysis, because they exhibited stably low expression (95-th percentile of \log_2 expression level less than 7).

Table 1 Number of patients in datasets

	Number of patients		
	Total number	Without recurrence and time of follow-up not less than 4 years	With recurrence within 3 years
GSE39582	519	245	124
GSE37892	129	48	32
GSE17536	144	58	31
GSE14333	225	85	41

Limitations of the study

All the data analyzed within the study have been acquired using Affymetrix Human Genome U133 Plus 2.0 Arrays and have not been confirmed by an independent method. Although generally reliable, problems with sensitivity, accuracy, and specificity have been reported for microarray results before [18]. Thus, specificity of microarray data may be compromised by non-specific hybridization that is a binding of not completely complimentary sequences to each other. However, it is worth noting that in a study [10] a gene signature identified using the same method based on similar microarrays (Affymetrix Human Genome U133A) successfully passed validation on a RNA-sequencing dataset. Another issue is possible discrepancies between expression of mRNAs and amounts of corresponding proteins due to differences in translation rates, half-lives of proteins, etc.

Irrespective of expression profiling technology, the method we used in this study for assessing prognostic power of gene sets and for finding small gene sets with the highest prognostic power has certain limitations. Lacking formal statistical background for associating p -values with the resulting prognosis reliability (expressed, e.g., in terms of AUC), it requires a large meta-set of samples (at least 120 samples in each group), and statistical significance can be assessed mainly using Monte Carlo resampling approaches. These limitations are not restrictive for this study.

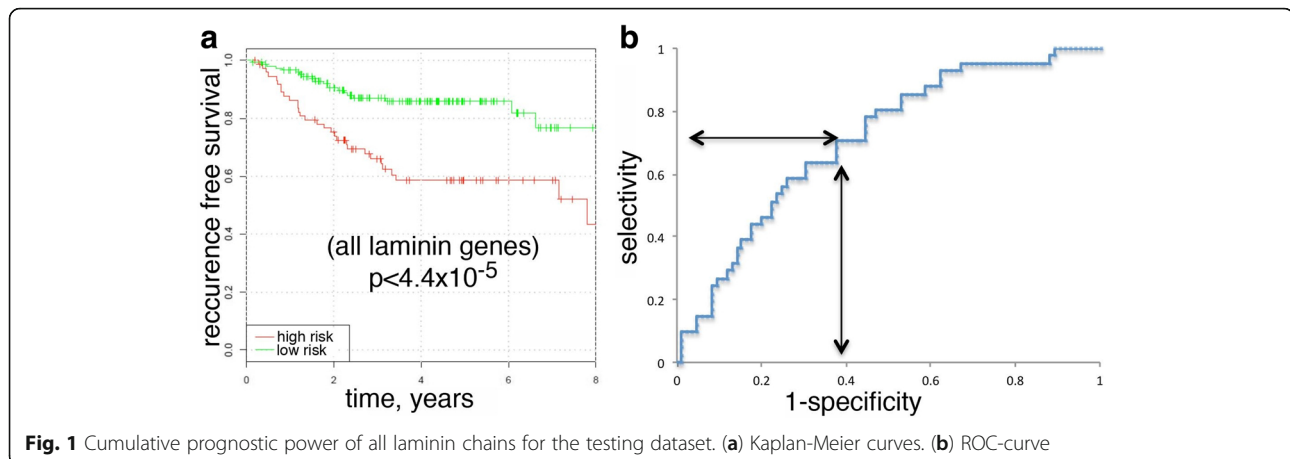
Statistics

Statistical significance of divergence of Kaplan-Meier survival curves was assessed using log-rank test. The presented p -values are two-tailed.

The statistical significance of the prognostic power of gene sets was assessed using Monte Carlo label permutation approach. Three permutation strategies were applied: labels permutation only in a training dataset, only in a testing dataset, and in both. Considering all permutations equiprobable, 10,000 permutations were generated for each strategy, and p -values were estimated based on AUC for the testing dataset in a standard way [19]. The highest (i.e., the most conservative) of three resulting values was reported.

Results

To assess prognostic power of the entire laminin family, we constructed a classifier based on all the laminin genes. The AUC value for the resulting classifier was equal to 0.710 for the testing dataset, and Monte Carlo label permutation approach confirmed the statistical significance of this AUC value (p -value $< 2.7 \times 10^{-3}$). Consistently, a divergence of Kaplan-Meier curves for this classifier (Fig. 1a) was also highly significant ($P = 4.4 \times 10^{-5}$). At the sensitivity level of 70% the specificity on the testing dataset noticeably exceeded 60% (Fig. 1b).



LAMA4, *LAMC1*, and *LAMC2* genes had the highest weights in the constructed classifier (Table 2). Higher expression of each of these genes was associated with an increase in the risk of colorectal cancer recurrence. Weights of the rest of the laminin genes were lower, but aggregated weight was comparable with that of the each of the three.

To determine whether the set of laminin genes could be reduced to *LAMA4*, *LAMC1*, and *LAMC2* or even to a smaller subset of the genes without a significant reduction of prognostic power, we performed the analysis for the subsets of the laminin genes, starting from single laminin chains (Fig. 2a-f). Surprisingly, only *LAMA4* gene and not *LAMC2* passed filtration threshold (see Materials and Methods). But even for *LAMA4*, the quality of the classifier was much lower than that for the classifier comprising all laminin genes. The AUC value of *LAMA4* classifier was equal to 0.62 for the testing dataset and a *p*-value

indicating statistical significance of Kaplan-Meier curves' divergence was equal to 0.036 (Fig. 2a and b).

Analysis of prognostic power of pairs of laminin genes revealed that any pair of the genes from *LAMA4*, *LAMC1*, and *LAMC2* provides a classifier, which successfully passes filtration stage. Classifiers based on (*LAMA4*, *LAMC1*), (*LAMA4*, *LAMC2*), and (*LAMC1*, *LAMC2*) had the AUC values equal to 0.643, 0.692, and 0.726, respectively, for the testing dataset and the *p*-value associated with Kaplan-Meier curves' divergence equal to 0.0057, 6.9×10^{-5} , and 0.0016, respectively. Therefore, prognostic power of (*LAMA4*, *LAMC2*) and (*LAMC1*, *LAMC2*) gene pairs was comparable with that of all laminin genes (Fig. 2c and d). Surprisingly, there were several other pairs containing other than *LAMA4*, *LAMC1*, and *LAMC2* genes with similar prognostic power in colorectal cancer. For instance, a classifier based on (*LAMA3*, *LAMC1*) had the AUC value 0.688 and the *P*-value 3.3×10^{-5} . Analysis of triples of laminin genes also revealed that other than (*LAMA4*, *LAMC1*, *LAMC2*) triples had similar prognostic power, for instance (*LAMA5*, *LAMC1*, *LAMC2*) and (*LAMA1*, *LAMA4*, *LAMC2*). In these triples, consistently with the classifier based on all laminin genes, higher expression of *LAMA1* and *LAMA5* was associated with lower risk of the recurrence.

Table 2 Weights of genes in the classifier based on all laminin chains. A positive value of the weight indicates that higher expression of the gene is associated with higher risk of recurrence. A negative weight indicates that higher expression of the gene is associated with lower risk of recurrence

Gene	Microarray identifiers	Weight
<i>LAMA1</i>	227048_at	-0.024
<i>LAMA2</i>	213519_s_at	-0.104
<i>LAMA3</i>	203726_s_at	0.010
<i>LAMA4</i>	202202_s_at	0.381
<i>LAMA5</i>	210150_s_at	-0.111
<i>LAMB1</i>	201505_at	-0.051
<i>LAMB2</i>	216264_s_at	0.011
<i>LAMB3</i>	209270_at	-0.049
<i>LAMC1</i>	200771_at	0.306
<i>LAMC2</i>	202267_at	0.324

Discussion

As expected, cumulative prognostic power of all laminin chains was higher than that of any single chain, but, interestingly, the prognostic power of all laminins was achieved already with some pairs and triples of the genes. *LAMA4*, *LAMC1*, and *LAMC2* genes had the highest weights in the classifier based on all laminin chains. Several independent groups have demonstrated prognostic value of *LAMC2* in colorectal cancer before [7, 8, 20]. To our knowledge, there is no published data on prognostic value of *LAMA4* and *LAMC1* in colorectal cancer, but $\alpha 4$ -containing laminins affect migration [21] and invasion [22] of various cancer

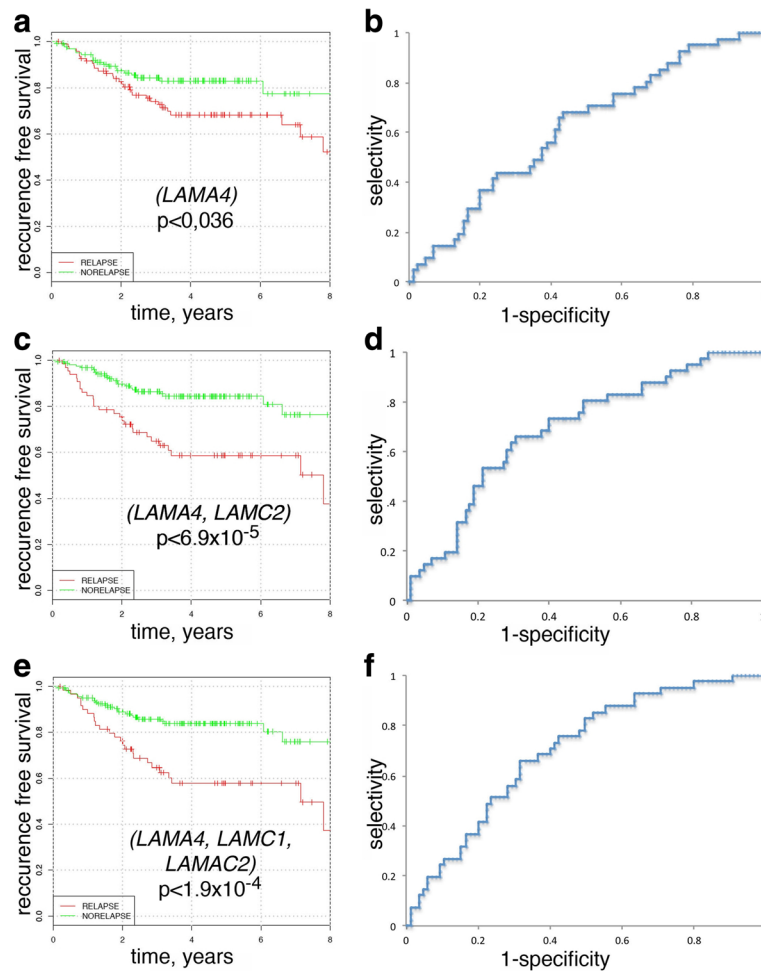


Fig. 2 Cumulative prognostic power of single, pairs, and triples of laminin chains for the testing dataset. Kaplan-Meier curves and ROC-curves for *LAMA4* (a and b, respectively); for (*LAMA4*, *LAMC2*) pair (c and d); and for (*LAMA4*, *LAMC1*, *LAMC2*) triple (e and f)

cells and expression of *LAMC1* is associated with cancer progression in uterine carcinoma [23] and glioblastoma [24].

Unlike previously identified laminin biomarkers, the prognostic gene signatures found in this study also suggest a specific molecular mechanism that might be involved in progression of colorectal cancer. Indeed, all the mentioned above predictive classifiers based on the triples of laminin genes might suggest an increased permeability of cancer BMs in patients with higher risk of colorectal cancer recurrence. High risk of recurrence was associated with increased expression of $\alpha 4$ or decreased expression of $\alpha 5$ laminin chains. Laminin polymer of the vascular BMs, which is an important component of the vascular wall, is primary made of $\alpha 4$ - and $\alpha 5$ -containing laminins [25]. It has been shown that leukocytes have a reduced ability to penetrate the vessel walls that are depleted of $\alpha 4$ -containing laminins [25, 26]. Since molecular mechanisms of diapedesis through the vessel wall for leukocytes and cancer cells are similar, we hypothesize that increased ratio of expression of $\alpha 4$ -

containing to $\alpha 5$ -containing laminins might lead to increased permeability of cancer BMs for cancer cells facilitating their detachment from the primary tumor. Therefore, BMs produced by cancer cells might be an important hindrance to dissemination of colorectal cancer. Although the hypothesis needs an experimental confirmation, it might shed new light on molecular mechanisms of colorectal cancer progression in patients.

Conclusions

We confirmed that laminins have a significant prognostic value in colorectal cancer. Prognostic power of any single laminin chain was lower than cumulative prognostic power of certain sets of laminin genes. Already some pairs and triples of laminin genes were sufficient to achieve the prognostic power of the entire laminin family. Detailed analysis of those pair and triples as well as genes weights in the classifiers based on laminin genes suggested that basement membranes produced by colorectal tumors might be an important hindrance to their own dissemination in patients.

Abbreviations

AUC: Area under ROC curve; BMs: Basement membranes; RMA: Robust Multi-array Average; ROC: Receiver operating characteristic; SVM: Support vector machine

Acknowledgements

We thank Dr. Maxim Yu. Shkurnikov and Dr. Timur R. Samatov for valuable comments and suggestions.

Funding

The study and the publication of this article were funded by the Russian Science Foundation (project 17-14-01338).

Availability of data and materials

All the datasets analyzed in this study were retrieved from the Gene Expression Omnibus which is a public-domain repository. Dataset accession numbers are specified in Materials and Methods section (subsection Microarray datasets).

About this supplement

This article has been published as part of BMC Medical Genomics Volume 11 Supplement 1, 2018: Selected articles from Belyaev Conference 2017: medical genomics. The full contents of the supplement are available online at <https://bmcmmedgenomics.biomedcentral.com/articles/supplements/volume-11-supplement-1>.

Authors' contributions

WVG, SR and AGT designed the study. WVG and AVG conducted in silico data analysis. DVM, SR and AGT interpreted the analysis results. WVG, DVM and SR wrote the manuscript. All authors read, revised and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Moscow State University, Leninskie Gory, 119991 Moscow, Russia. ²SR Bioclinicum, Ugreshskaya str 2/85, 115088 Moscow, Russia. ³Tauber Bioinformatics Research Center, University of Haifa, Mount Carmel, 3498838 Haifa, Israel. ⁴Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Scheelesväg 2, 17177 Stockholm, SE, Sweden. ⁵P. Hertsen Moscow Oncology Research Institute, National Center of Medical Radiological Research, Second Botkinsky Lane 3, 125284 Moscow, Russia.

Published: 13 February 2018

References

- Aumailley M, Bruckner-Tuderman L, Carter WG, Deutzmann R, Edgar D, Ekblom P, Engel J, Engvall E, Hohenester E, Jones JC, et al. A simplified laminin nomenclature. *Matrix Biol.* 2005;24(5):326–32.
- Yurchenco PD. Basement membranes: cell scaffoldings and signaling platforms. *Cold Spring Harb Perspect Biol.* 2011;3(2)
- Jayadev R, Sherwood DR. Basement membranes. *Curr Biol.* 2017;27(6):R207–11.
- Qin Y, Rodin S, Simonson OE, Hollande F. Laminins and cancer stem cells: partners in crime? *Semin Cancer Biol.* 2017;45:3–12.
- Saito N, Kameoka S. Serum laminin is an independent prognostic factor in colorectal cancer. *Int J Color Dis.* 2005;20(3):238–44.
- Fukazawa S, Shinto E, Tsuda H, Ueno H, Shikina A, Kajiwara Y, Yamamoto J, Hase K. Laminin beta3 expression as a prognostic factor and a predictive marker of chemoresistance in colorectal cancer. *Jpn J Clin Oncol.* 2015;45(6):533–40.
- Huang D, Du C, Ji D, Xi J, Gu J. Overexpression of LAMC2 predicts poor prognosis in colorectal cancer patients and promotes cancer cell proliferation, migration, and invasion. *Tumour Biol.* 2017;39(6):1010428317705849.
- Shinto E, Tsuda H, Ueno H, Hashiguchi Y, Hase K, Tamai S, Mochizuki H, Inazawa J, Matsubara O. Prognostic implication of laminin-5 gamma 2 chain expression in the invasive front of colorectal cancers, disclosed by area-specific four-point tissue microarrays. *Lab Invest.* 2005;85(2):257–66.
- Sordat I, Bosman FT, Dorta G, Rousselle P, Aberdam D, Blum AL, Sordat B. Differential expression of laminin-5 subunits and integrin receptors in human colorectal neoplasia. *J Pathol.* 1998;185(1):44–52.
- Galatenko VV, Shkurnikov MY, Samatov TR, Galatenko AV, Mityakina IA, Kaprin AD, Schumacher U, Tonevitsky AG. Highly informative marker sets consisting of genes with low individual degree of differential expression. *Sci Rep.* 2015;5:14967.
- Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20(3):273–97.
- Marisa L, de Reynies A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-Grimaldi MC, Schiappa R, Guenot D, Ayadi M, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* 2013;10(5):e1001453.
- Laibe S, Lagarde A, Ferrari A, Monges G, Birnbaum S, Project COL. A seven-gene signature aggregates a subgroup of stage II colon cancers with stage III. *OMICS.* 2012;16(10):560–5.
- Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, Lu P, Johnson JC, Schmidt C, Bailey CE, et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology.* 2010;138(3):958–68.
- Freeman TJ, Smith JJ, Chen X, Washington MK, Roland JT, Means AL, Eschrich SA, Yeatman TJ, Deane NG, Beauchamp RD. Smad4-mediated signaling inhibits intestinal neoplasia by inhibiting expression of beta-catenin. *Gastroenterology.* 2012;142(3):562–71. e562
- Jorissen RN, Gibbs P, Christie M, Prakash S, Lipton L, Desai J, Kerr D, Aaltonen LA, Arango D, Kruhoffer M, et al. Metastasis-associated gene expression changes predict poor outcomes in patients with dukes stage B and C colorectal cancer. *Clin Cancer Res.* 2009;15(24):7642–51.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics.* 2003;4(2):249–64.
- Draghici S, Khatri P, Eklund AC, Szallasi Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet.* 2006;22(2):101–9.
- Ewens WJ. On estimating P values by the Monte Carlo method. *Am J Hum Genet.* 2003;72(2):496–8.
- Lenander C, Habermann JK, Ost A, Nilsson B, Schimmelpenninck H, Tryggvason K, Auer G. Laminin-5 gamma 2 chain expression correlates with unfavorable prognosis in colon carcinomas. *Anal Cell Pathol.* 2001;22(4):201–9.
- Ishikawa T, Wondimu Z, Oikawa Y, Gentilcore G, Kiessling R, Egyhazi Brage S, Hansson J, Patarroyo M. Laminins 411 and 421 differentially promote tumor cell migration via alpha6beta1 integrin and MCAM (CD146). *Matrix Biol.* 2014;38:69–83.
- Nagato S, Nakagawa K, Harada H, Kohno S, Fujiwara H, Sekiguchi K, Ohue S, Iwata S, Ohnishi T. Downregulation of laminin alpha4 chain expression inhibits glioma invasion in vitro and in vivo. *Int J Cancer.* 2005;117(1):41–50.
- Kashima H, Wu RC, Wang Y, Sinno AK, Miyamoto T, Shiozawa T, Wang TL, Fader AN, Shih le M. Laminin C1 expression by uterine carcinoma cells is associated with tumor progression. *Gynecol Oncol.* 2015;139(2):338–44.
- Huang SX, Zhao ZY, Weng GH, He XY, Wu CJ, Fu CY, Sui ZY, Zhong XM, Liu T. The correlation of microRNA-181a and target genes with poor prognosis of glioblastoma patients. *Int J Oncol.* 2016;49(1):217–24.
- Sixt M, Engelhardt B, Pausch F, Hallmann R, Wendler O, Sorokin LM. Endothelial cell laminin isoforms, laminins 8 and 10, play decisive roles in T cell recruitment across the blood-brain barrier in experimental autoimmune encephalomyelitis. *J Cell Biol.* 2001;153(5):933–46.
- Kenne E, Soehnlein O, Genove G, Rotzius P, Eriksson EE, Lindbom L. Immune cell recruitment to inflammatory foci is impaired in mice deficient in basement membrane protein laminin alpha4. *J Leukoc Biol.* 2010;88(3):523–8.