Inherited variation in pattern recognition receptors and cancer: dangerous liaisons?

Anton G Kutikhin
Arseniy E Yuzhalin
Department of Epidemiology, Kemerovo State Medical Academy, Kemerovo, Russian Federation

Abstract: The group of pattern recognition receptors includes families of Toll-like receptors, NOD-like receptors, C-type lectin receptors, and RIG-I-like receptors. They are key sensors for a number of infectious agents, some of which are carcinogenic, and they launch an immune response against them. Inherited structural variation in genes encoding these receptors and proteins of their signaling pathways may affect their function, modulating cancer risk and features of cancer progression. Relevant malignancies, valuable gene polymorphisms, prime questions about future directions, and answers to these questions are analyzed in this review. It is possible to suggest that polymorphisms of genes encoding pattern recognition receptors and proteins of their signaling pathways may be associated with almost all cancer types, particularly with those in which carcinogenic infectious agents are responsible for the substantial share of cases (namely gastric cancer, colorectal cancer, liver cancer, cervical cancer, and nasopharyngeal carcinoma). The concept of selection of polymorphisms for further oncogenomic investigation, based on a combination of results from basic and epidemiological studies, is proposed.

Keywords: pattern recognition receptors, Toll-like receptors, NOD-like receptors, C-type lectin receptors, cancer, gene polymorphisms

Introduction
What are pattern recognition receptors?
The group of pattern recognition receptors (PRRs) includes families of Toll-like receptors (TLRs), NOD-like receptors, C-type lectin receptors (CLRs), and RIG-I-like receptors (RLRs). A summary of the most modern conceptual data about members of these families and their structure and functions can be obtained from recent comprehensive papers of Kawai and Akira,1 Elinav et al,2 Osorio and Reis e Sousa,3 and Loo and Gale.4 Receptors constituting these families are united by two general features. Firstly, they directly recognize common antigen determinants of virtually all classes of pathogens (so-called pathogen-associated molecular patterns) and initiate immune response against them via specific intracellular signaling pathways.1–4 Secondly, they recognize endogenous ligands (since they are usually released during cell stress, they are called damage-associated molecular patterns), and, consequently, PRR-mediated immune response can be activated without the influence of infectious agents.1–4 Therefore, PRRs may also initiate development of aseptic inflammation caused by physical factors such as mechanical pressure, thermal damage, ionizing and nonionizing radiation, or chemical factors (eg, acidic damage, alkaline damage, exposure to chemical war gases, croton oil or turpentine, exposure to allergens, liberation of toxic
substances during tumor disintegration, aseptic necrosis, internal bleeding, hemolysis, and autoimmune processes.\textsuperscript{1–4} It may promote further progression of inflammation or, on the contrary, prevent hazardous infectious complications (the combination of these two effects may also be true).\textsuperscript{1–4} The final outcome of PRR working is enhanced production of many proinflammatory cytokines participating in plenty of the immune system’s processes.\textsuperscript{1–4} Expression of PRRs on different levels (transcriptomic or proteomic) was detected in a lot of cells and organs,\textsuperscript{1–4} providing evidence that these receptors control many elements of the complex machinery of the human immune system: they allow epithelium and endothelium to defend against infectious agents on their own, they mediate the activation of adaptive immune response by antigen-presenting cells and T-helpers, they stimulate expression of cell adhesion molecules for leukocyte rolling and for other processes of inflammation development, and, finally, they contribute to phagocytosis efficacy.\textsuperscript{5} As a consequence of all mentioned above, PRRs play a key role in the realization of innate and adaptive immune response. In addition, many PRRs have a number of other vital functions apart from participation in immune response realization: they may regulate various aspects of cell proliferation, survival, apoptosis, autophagy, reactive oxygen species generation, pyroptosis, angiogenesis, and, consequently, of tissue remodeling and repair.\textsuperscript{6–9}

The fundamental characteristics and diversity of PRR functions have led to amazingly rapid research in this field, and such investigations are very promising for medicine as the immune system plays a key role in the vast majority, if not all, human diseases, and the process of discovering new aspects of immune system functioning is rapidly ongoing. There is a plethora of papers analyzing the significance of PRRs in various diseases. One of the most actively explored fields in PRR biology is their role in cancer etiopathogenesis. Not surprisingly, it is (as well as tumor immunology in general) a “hot spot” in cancer biology as well.

The role of pattern recognition receptors in cancer development

Since PRRs mediate the immune response induced by many immunoadjuvants,\textsuperscript{10,11} and many of them regulate the immune response against potentially carcinogenic infectious agents\textsuperscript{12,13} (\textit{Helicobacter pylori},\textsuperscript{14–17} Epstein–Barr virus (EBV),\textsuperscript{18,19} human papillomavirus,\textsuperscript{18,19} human herpesvirus-8/ Kaposi sarcoma-associated herpesvirus,\textsuperscript{18,19} \textit{Mycobacterium tuberculosis},\textsuperscript{15–17,20} \textit{Streptococcus pneumoniae},\textsuperscript{15–17,21} enteropathogenic \textit{Escherichia coli},\textsuperscript{15–17,22} \textit{Shigella flexneri},\textsuperscript{15–17,23} \textit{Salmonella typhimurium},\textsuperscript{15–17,24} \textit{Borrelia burgdorferi},\textsuperscript{15–17,25} \textit{Chlamydomphila pneumoniae},\textsuperscript{15–17} \textit{Chlamydia psittaci},\textsuperscript{15–17} and \textit{Campylobacter jejuni},\textsuperscript{15–17,27}), it seems to be possible to stimulate antitumor immunity through their enhanced activation.\textsuperscript{28,29} This hypothesis, originally developed for TLRs, should be also true for all PRRs as well.\textsuperscript{28,29} According to this suggestion, reinforced PRR activation may protect from infectious agents and prevent, inhibit, or block carcinogenesis whilst disrupted functioning of these PRRs may allow infectious agents or tumor cells to avoid recognition by the immune system and, consequently, not be eliminated.\textsuperscript{28,29} At the same time, such PRR activation may promote carcinogenesis, creating a proinflammatory microenvironment (via action of respective cytokines) that is favorable for tumor progression and chemoresistance development.\textsuperscript{30} It may also result in immunosuppression caused by chronic inflammation.\textsuperscript{28} Chronic inflammation may promote the development of cervical, endometrial, ovarian, breast, prostate, testicular, nasopharyngeal, lung, esophageal, gastric, colorectal, liver, pancreatic, gallbladder, kidney, bladder, lymphatic malignancies, and feasibly several other cancer types.\textsuperscript{11,31} In this case, on the contrary, lower PRR activity should minimize effects of chronic inflammation such as enhancement of cancer initiation and promotion/progression and, consequently, decrease probability of tumor development.\textsuperscript{30} So, the situation resembles a double-edged sword. The ideal variant, possibly, is the “golden mean” – the balance between low and high PRR activity. This hypothesis, initially developed for PRRs,\textsuperscript{29} may also be successfully projected on PRR intracellular signaling pathways; if their elements are overexpressed/constantly activated, it may lead to similar consequences as enhanced PRR activation. On the other hand, if members of PRR pathways are underexpressed/inactivated/unable to do their work, it may result in the same effects that arise after decreased PRR activity, and the analogical “golden mean” in functioning of all genes encoding proteins constituting PRR signaling pathways will be the optimal variant.

Structural genomic variation

The completion of the Human Genome Project and widespread distribution of genotyping technologies have led to an enormous number of studies devoted to the association of inherited gene polymorphisms with various diseases. Single nucleotide polymorphisms (SNPs) may result in amino acid substitutions altering protein function or splicing, and they can also change the structure of enhancer sequences during splicing\textsuperscript{32} or affect
messenger ribonucleic acid (RNA) stability. SNPs may alter transcription factor binding motifs, changing the efficacy of enhancer or repressor elements, and can alter the structure of translation initiation codons that may lead to downregulation of wild-type transcript. Gene polymorphisms located in leucine-rich repeats constituting ectodomain of PRRs may affect the ability of the receptor to bind pathogens they normally recognize. SNPs in transmembrane domain can lead to defects of intracellular receptor transport that do not allow location of a receptor on the membrane, and, finally, polymorphisms in the internal domain may result in altered interaction with adaptor proteins or in disrupted dimerization. So, inherited SNPs of genes encoding PRRs may alter PRR expression and activity, modulating cancer risk and, possibly, influencing various features of cancer progression. The same statement should be true for genes encoding proteins of PRR signaling pathways.

Based on the plethora of fundamental and epidemiological studies carried out, it is possible to specify two fundamental mechanisms for modulation of cancer risk by polymorphisms of genes encoding PRRs and proteins of PRR pathways. The first mechanism is impairment of the immune response to certain pathogens (it can be bacteria, viruses, fungi, protozoan, and helminths), which increases the risk of potentially carcinogenic infection and promotes its development along with further chronic persistence. The second mechanism is an increase in production of proinflammatory cytokines after binding of the ligand (exogenous or endogenous), which creates a condition of carcinogenic chronic inflammation.

**Relevant malignancies: the first dimension of investigation**

There is a variety of cancer types definitely or possibly having infectious etiology that can be associated with inherited alterations in genes encoding PRRs and proteins of PRR signaling pathways:

- esophageal cancer (variation in immune response to pathogens infecting esophagus)
- gastric cancer (on the basis of modulation of immune response to *H. pylori*)
- cancer of the small bowel (modulation of immune response to *C. jejuni*)
- colorectal cancer (alteration of immune response to many, mostly undefined, infectious agents inhabiting the colon and rectum)
- liver cancer (variation in immune response to hepatitis B virus, hepatitis C virus, *Helicobacter hepaticus*, or liver flukes)
- gallbladder cancer (modulation of immune response to infectious agents found in bile)
- pancreatic cancer (alteration of immune response to pathogens inhabiting the pancreas)
- endometrial cancer (modification of immune response to several kinds of infectious agents colonizing the endometrium)
- cervical cancer (alteration of immune response to human papillomavirus and some infectious agents colonizing the cervix)
- ovarian cancer (variation in immune response to *C. trachomatis*)
- breast cancer (modulation of immune response to some viruses infecting the breast)
- prostate cancer (variation in immune response to *Propionibacterium acnes* and other uncertain pathogens found in prostate tissue)
- testicular cancer (modification of immune response to EBV)
- kidney cancer (variation in immune response to bacteria and viruses infecting the kidneys)
- bladder cancer (modulation of immune response to certain viruses or *Schistosoma* spp.)
- nasopharyngeal carcinoma (alteration of immune response to EBV)
- lung cancer (variation in immune response to *M. tuberculosis*, *S. pneumoniae*, *C. pneumoniae*, and, possibly, to other infectious agents causing chronic inflammatory lung diseases)
- lymphoma (modification of immune response to EBV and many other infectious agents such as *B. burgdorferi* or *H. pylori*)

**Selection of valuable polymorphisms: the second dimension of investigation**

It is important to remember that there are two main components determining the importance of the SNP in programs of cancer prevention based on genomic risk markers: the odds ratio value between cases and controls (as in the whole population and subgroups) and the prevalence of the polymorphism in the population, and they both may vary in different geographic regions. It is desirable to develop not one general program, but a number of individual programs for different countries/populations. At the moment, it is
possible only to recommend a list of polymorphisms for further investigation since only a small number of studies with perfect design were carried out. The list of relevant polymorphisms that can be admitted as the most promising for further oncogenomic investigations may be created according to the following rules.

A gene polymorphism may be included into the short list for further oncogenomic studies if:

• The SNP leads to substantial functional consequences on the molecular level (for instance, it strongly affects transcription, splicing, translation, stability and transport of pre-messenger RNA, messenger RNA, noncoding RNA, or protein encoding by the gene, or it noticeably influences signaling of synthesized protein)

• It is associated with the risk of cancer in conducted studies

• It has any functional consequences on the molecular level and it is strongly (threshold odds ratio value may be individual for each cancer type) associated with conditions that significantly increase the risk of cancer.

A gene polymorphism can be also included into the extended list if:

• It is characterized by more subtle functional alterations in the gene that, however, still result in qualitative or quantitative alterations of the encoding protein (or noncoding RNA)

• It is associated only with conditions that substantially increase the risk of cancer (ie, not associated with the risk of cancer).

In concordance with this conception, the following SNPs of genes encoding PRRs and proteins of PRR signaling pathways may be accepted as the most valuable for further oncogenomic investigations based on the analysis of relevant published literature:63–65

• TLR1-TLR6-TLR10 gene cluster: rs10008492, rs4833103, rs5743815, rs11466657

• TLR2 gene: rs3804100, rs4696480, −196 to −174 del (Delta22), GT-microsatellite polymorphism

• TLR4 gene: rs4986790, rs4986791, rs16906079, rs11536891, rs8737384, rs1927911, rs10759932, rs10116253, rs11536889, rs11536858

• TLR9 gene: rs5743836, rs352140

• TIRAP/MAL gene: rs8177400, rs8177399, rs8177374, rs7932766

• MyD88 gene: rs1319438, rs199396

• IRAKI gene: rs1059703, rs3027898, rs10127175

• TRAF3 gene: rs7143468, rs12147254, rs11160707

• TOLLIP gene: rs5743867

• IRF3 gene: rs7251

• IRF5 gene: rs2004640, rs2280714, rs10954213, 5 bp indel (CGGGG) polymorphism

• NOD1 gene: rs2075820, ND(1) + 32656

• NOD2 gene: rs2066842, rs2066844, rs2066845, rs2006847

• MRC1 gene: rs1926736, rs2478577, rs2437257, rs691005

• CD209 gene: rs2287886, rs735239, rs735240, rs4804803

• CLEC7A gene: rs16910526

• RIG-I gene: rs36055726, rs11795404, rs10813831.

Although gene polymorphisms of genes encoding RLRs, CLRs, and specific proteins of their signaling pathways are investigated relatively less than SNPs of TLRs and NOD-like receptors, it is possible to conclude that since they recognize bacterial, viral, fungal, protozoan, and helminth pathogen-associated molecular patterns as TLRs and NOD-like receptors, inherited structural variation in them may influence cancer risk and progression as well. For instance, some human CLRs (MRC1, CD207, LY75, CD209, CLEC7A, CLEC1B, CLEC6A, CLEC4E, CLEC4A) recognize ligands of potentially carcinogenic infectious agents,12,13 such as M. tuberculosis,66 S. pneumoniae,66 Klebsiella pneumoniae,68 human immunodeficiency virus-1,69 cytomegalovirus,70 Candida albicans,71 C. neoformans,72 Pneumocystis carinii,73 Paracoccidioides brasiliensis,74 Histoplasma capsulatum,74 and Schistosoma mansoni.75 Many polymorphisms of genes encoding these receptors may alter immune response to indicated ligands, possibly, modulating etiopathogenesis of certain cancer types such as lung cancer (M. tuberculosis, S. pneumoniae, and K. pneumoniae),75 glioblastoma (cytomegalovirus),76 oral cancer (fungi),77 colorectal cancer, hepatocellular carcinoma, prostate cancer, or cervical cancer (S. mansoni).78 One RLR, RIG-I, also recognizes ligands of hepatitis C virus and EBV,78 and thus structural inherited variation in this receptor may alter risk of hepatocellular carcinoma, nasopharyngeal carcinoma, and lymphoma.

There are certain disparities in different population studies investigating the association of polymorphisms of genes encoding PRRs and proteins of their signaling pathways with various aspects of cancer development.63–65 General reasons for these discrepancies may include confounding host, bacterial, or environmental factors in different ethnicities modulating penetrance of variant allele and affecting risk of conditions increasing cancer risk (such as autoimmune diseases, precancerous gastric lesions, tuberculosis,
and recurrent pneumonia), different bacterial impact in etiology of such conditions in different populations (that will be reflected in different features of PRR-mediated immune response because of specific PRR-ligand interaction), differences in sample size, differences in age/gender/ body mass index/ethnicity/tumor, node, metastasis stage/ other clinicopathological characteristics between study samples, differences in prevalence of infectious agent (eg, *Helicobacter pylori* or EBV) in case and control groups, and differences in diagnostics, stratification, genotyping methods, and chance. In addition, certain studies in which negative results were obtained could never have been published (so-called “file drawer effect”), which may create a significant bias and distort a picture that cannot be observed at the moment. Unfortunately, although some genome-wide association studies relevant to the discussed problem were performed, it is usually not possible to compare them with non-genome-wide association studies on the same cancer type since there are no non-genome-wide association studies investigating association of the same SNPs with similar malignancies. It may be feasible in the future when the number of studies devoted to this issue will be enough for correct comparative analysis.

**Future directions**

The most intriguing aspects of the problem of the association of inherited structural variation in genes encoding PRRs and proteins of PRR signaling pathways with features of cancer development are:

- Are SNPs in genes encoding PRRs or proteins of PRR signaling pathways associated with features of cancer progression or only with cancer risk? Existing studies show controversial results, and most of the results suggest that there is no, or weak, correlation between such polymorphisms and peculiarities of cancer progression.
- Are polymorphisms of genes encoding CLRs, RLRs, or specific proteins of their signaling pathways connected with risk or progression of cancer? If yes, would it be appropriate to include them in the list of polymorphisms used in programs of cancer risk determination and further cancer prevention? As shown above, there is some premise to the thought that these SNPs may be associated with cancer risk. Further fundamental and population studies are necessary to answer to this question.
- Do polymorphisms of genes encoding PRRs or proteins of PRR signaling pathways (particularly TLRs and TLR pathway) correlate with altered prostate cancer risk or progression? Almost all large studies devoted to this issue showed that there is no association between inherited variation in indicated genes and features of prostate cancer development.
- Are polymorphisms of genes of PRR signaling pathways associated with cancer risk or progression to the same extent as polymorphisms of genes encoding PRRs? It is logical that if SNP of gene encoding specific PRR is associated with risk or progression features of certain malignancies, polymorphisms in genes encoding specific signaling molecules constituting pathways of these receptors should correlate with similar neoplasms, if they have substantial functional consequences on the molecular level.
- How do polymorphisms of genes encoding PRRs and proteins of PRR signaling pathways interact with each other in relation to determination of cancer risk and progression? Particularly, how do SNPs of positive and negative regulators of PRR activity (especially, micro RNA) influence cancer risk or progression if they are inherited together? Answers to these questions remain elusive at the present time and should be obtained from fundamental and population studies in the future.
- Which SNPs of genes encoding PRRs and proteins of PRR pathways have independent significance, and which are just in the linkage disequilibrium? Knowledge of this may help in creating the list of polymorphisms useful in programs of cancer risk determination and further prevention.
- Which SNPs of genes encoding PRRs and proteins of PRR pathways should be included in such a list? Which of them have universal effect for each cancer type, and which influence risk and progression of one cancer type but have no effect in relation to another malignancy? Differences in association of the same SNP with different malignancies should be explained by features of specific pathogen-associated molecular pattern–PRR interaction (probably certain characteristics of ligand binding), or, possibly, on peculiarities of damage-associated molecular pattern–PRR interaction. Lists of prospective SNPs for further oncogenic investigations may be created according to the concept suggested above.
- How do SNPs of genes encoding PRRs and proteins of PRR pathways affect cancer risk or progression in different populations and subgroups of such populations? How can this information be adjusted for application in
the creation of programs of cancer risk determination and further prevention? Only large, comprehensive, well-designed population studies may provide an answer to these questions.

- Do polymorphisms of genes encoding PRRs and proteins of PRR pathways influence cancer risk only through increase in risk of chronic inflammatory conditions, or can they affect it through other mechanisms as well? How can this information be used in programs of cancer risk determination and further prevention? To answer these questions, control groups in population studies should include not only healthy controls, but also controls with chronic inflammatory conditions predisposed to investigating cancer type.

- Which infectious agents recognized by various PRRs are carcinogenic, and which are not? It may help to define cancer types associated with SNPs of genes encoding specific PRRs and proteins constituting PRR signaling pathways. Fundamental studies devoted to the investigation of infectious agent–PRR interactions, to the investigation of carcinogenicity of known infectious agents and to the discovery of new, possibly carcinogenic, infectious agents should answer this question.

No doubt, determination of the role of SNPs in genes encoding PRRs and proteins of PRR signaling pathways in fields of tumor immunology and molecular epidemiology of cancer may open new pages in cancer biology and cancer prevention.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**

