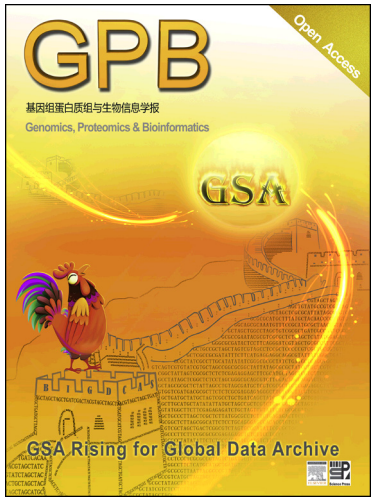
Accepted Manuscript



Omics Approaches to Identify Potential Biomarkers of Inflammatory Diseases in the Focal Adhesion Complex

Johanne Brooks, Alastair Watson, Tamas Korcsmaros

PII: S1672-0229(17)30041-4
DOI: http://dx.doi.org/10.1016/j.gpb.2016.12.003
Reference: GPB 240

To appear in: *Genomics, Proteomics & Bioinformatics*

Received Date: 26 July 2016
Revised Date: 13 December 2016
Accepted Date: 21 December 2016

Please cite this article as: J. Brooks, A. Watson, T. Korcsmaros, Omics Approaches to Identify Potential Biomarkers of Inflammatory Diseases in the Focal Adhesion Complex, *Genomics, Proteomics & Bioinformatics* (2017), doi: http://dx.doi.org/10.1016/j.gpb.2016.12.003

Omics Approaches to Identify Potential Biomarkers of Inflammatory Diseases in the Focal Adhesion Complex

Johanne Brooks[1,2,3,a], Alastair Watson[1,2,3,b], Tamas Korcsmaros[1,4,*,c]

[1] *Gut Health and Food Safety Institute Strategic Programme, Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, United Kingdom*

[2] *Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom*

[3] *Gastroenterology Department, Norfolk and Norwich University Hospital, Norwich NR4 7UY, United Kingdom*

[4] *Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, United Kingdom*

[*]Corresponding author.

E-mail: Tamas.Korcsmaros@earlham.ac.uk (Korcsmaros T).

Running title: *Brooks J et al / Omics for Biomarkers in the Focal Adhesion Complex*

[a] ORCID: 0000-0001-5967-7225.

[b] ORCID: 0000-0003-3326-0426.

[c] ORCID: 0000-0003-1717-996X.

Total count of words: 5928

Total count of figures: 3

Total count of tables: 1

Abstract

Inflammatory diseases such as inflammatory bowel disease (IBD) require recurrent invasive tests, including blood tests, radiology, and endoscopic evaluation both to diagnose and assess disease activity, and to determine optimal therapeutic strategies. Simple 'bedside' biomarkers could be used in all phases of patient management to avoid unnecessary investigation and guide further management. The focal adhesion complex (FAC) has been implicated in the pathogenesis of multiple inflammatory diseases, including IBD, rheumatoid arthritis, and multiple sclerosis. Utilizing omics technologies has proven to be an efficient approach to identify biomarkers from within the FAC in the field of cancer medicine. Predictive biomarkers are paving the way for the success of precision medicine for cancer patients, but inflammatory diseases have lagged behind in this respect. This review explores the current status of biomarker prediction for inflammatory diseases from within the FAC using omics technologies and highlights the benefits of future potential biomarker identification approaches.

KEYWORDS: Focal adhesion complex; Biomarkers; Inflammation; Omics; Systems biology

Introduction

Disease biomarkers have the potential to be medically valuable at all stages of the disease process from diagnosis, identification of disease subtypes, and prognosis to therapeutic adjustment. Inflammatory bowel disease (IBD) is an exemplar of a chronic, complex inflammatory disease. IBD has two major subtypes, ulcerative colitis (UC) and Crohn's disease, which have different clinical courses and management strategies with a wide phenotypic variability among patients. **Figure 1** highlights the points at which biomarkers have potential use in IBD.

Biomarkers need to be specific, stable, and consistent across multiple platforms of testing in order to be used as a clinical application. This raises challenges associated with biomarker identification in IBD, as with any complex inflammatory condition, partly due to our limited understanding of the pathogenesis of these diseases and poor appreciation of the difference between what is healthy and what is a disease process. Hypothesis-driven biomarker discovery via traditional one protein–one metabolite or one cell analysis from cellular disease models or tissues compared between control and disease samples is laborious. Such an approach is also limited by the fact that gene expression and signalling of tissues depends on the context and their native environments [1]. For this reason, very few biomarkers make it to clinical practice [2]. Further challenges posed by complex diseases are that they often need to be stratified into sub-phenotypes via patients' genetic features, which need to be taken into account, making identification of a broad generalizable biomarker difficult [3]. High throughput, hypothesis-free techniques are required for biomarker discovery. With the advent of high-throughput omics technologies and advances in computational biology, researchers are now able to generate, analyze, and interpret a variety of datasets and apply them on biomarker discovery at a scale, which were previously impossible (**Figure 2**). One of the cellular signal transduction pathways supplying candidate biomarkers that have become prominent through the use of omics technologies and computational biology, certainly for the cancer field, is the focal adhesion complex (FAC).

FACs are dynamic, large protein assemblies that mechanically link and transduce signals from the extracellular matrix to the intracellular milieu via integrins [4] or other receptor modules such as cluster of differentiation 47 (CD47). The complex consists of core structural proteins such as paxillin, talin, actinin, and vinculin, with dynamic signalling proteins including protein kinases, phosphatases, small guanosine triphosphatases (GTPases) with regulatory molecules, and adapter molecules that mediate core protein–protein interactions

(**Table 1**). The 'adhesome' network contains 156 components with 690 interactions between them [26], highlighting the complexity of the focal adhesion function.

The focal adhesion function is both mechanical and responsive. It is mechanical in terms of anchoring the cell to the extracellular matrix via binding of integrins to their extracellular ligands and to the actin cytoskeleton to modify the physical and topographical characteristics of the cell. This has direct implications for wound healing as well as invasion and the metastatic nature of the cancer cell. The responsive function of the FAC is diverse and multi-layered. Depending on the initiating signal, FAC can be involved in regulating inflammatory gene expression via signal transduction pathways such as interleukin 1 (IL-1) signalling [27,28] or regulating calcium fluxes via phosphatidyl inositol signalling [29], which impact on inflammatory cascades. Many molecules in the FAC are involved in downstream signalling pathways, for instance, the MAPK/ERK pathway [30], AKT1 [22], and Wnt signalling [31,32]. In this way, pathways impacted by the FAC are as varied as apoptosis [21], production of cellular protrusions [33], cell cycle progression [34], and cell proliferation [35].

The number of publications listed in PubMed involving FAC ('focal adhesion complex') has had a 5-fold increase from 141 published in 1996 to 709 published in 2015. The role of FAC in cancer has been a consistent focus of approximately 44% of publications over the past 20 years (**Figure 3**). Given the critical roles that focal adhesions play in regulating cell structure, proliferation, survival, migration, and invasion, it is not surprising that this makes the complex a prime target for biomarker candidacy and drug targeting in cancer, which is reflected in the overrepresentation of papers with the terms 'cancer', 'focal adhesion', and 'biomarker' from a cohort of 'focal adhesion' and 'biomarker' publication subset.

Of the publications identified using the Medical Subject Headings (MeSH) terms 'cancer', 'focal adhesion', and also adding 'biomarker', 39 out of 745 used bioinformatics approaches for biomarker identification. It is of note that all these 39 studies were published after 2007.

The role of FAC in the pathobiology of inflammatory diseases such as IBD or rheumatoid arthritis (RA) has been less well exploited for biomarker discovery. However, the role of FAC in inflammatory diseases can be well illustrated in UC. UC is a relapsing-remitting disease which causes ulceration of the lining of the large bowel and is thought to be a disease of the epithelial barrier [36]. The epithelial barrier is an immuno-mechanical barrier consisting of mucous layers, intestinal epithelial cells, and closely-residing immune cell populations. The mechanical barrier is provided in part by the enterocytes joined by

intercellular junctions, of which the tight junction is a major component. May et al. [37] identified that activation of focal adhesion kinase (FAK) is necessary for maintaining and repairing the epithelial barrier in cell culture via tight junctions. This was further examined by Khan et al. [38] in both T84 cell lines and surgical specimens from IBD patients. They demonstrated that activation of M1 muscarinic acetylcholine receptor augmented the recovery of epithelial barrier function via phosphorylation of FAK. Further evidence for the role of FAK in maintaining intestinal epithelial barrier function in the presence of pathogenic factors was highlighted by Guo and colleagues [39]. Utilizing intestinal epithelial cell cultures, they identified that gut-derived bacterial lipopolysaccharide induced tight junction permeability via the FAK/myeloid differentiation primary response gene 88 (MyD88)/IL1 receptor pathway. GTPases such as Rac1 [40] and tyrosine phosphatase members of FAC have a role in regulation of the NACHT, LRR and PYD domains-containing protein 3 (NLRP3; also known as cryopyrin) inflammasome [41], which mediates the release of IL-1 and IL-18 from cells. IL-18 signalling drives the breakdown of barrier integrity in murine models of UC [42]. Further evidence of FAC involvement in inflammasome activation was provided by Thinwa et al. [43] who demonstrated that the initial signal for intestinal cell inflammasome activation in pathogen recognition is via integrins. It is interesting to note that *NLRP3* was identified as a candidate gene for susceptibility of Crohn's disease [44], whereas IL-1 has been put forward as a faecal marker of inflammation in UC [45].

The evidence described above has been hypothesis-driven, utilizing mainly cellular models to describe a pathogenic system. In this review we will consider the literature field of FAC in inflammatory diseases focusing on those utilizing a systems medicine approach, where omics data and computational biology are combined for potential biomarker identification.

In the last two decades, omics technologies have made a great impact on medical research, turning biological research into a data-intensive science [46]. These high-throughput methodologies are now routinely used to provide a top-down approach in understanding biological systems. The power of omics approaches in systems medicine is due to their ability to detect context (*e.g.*, cell, disease, or treatment) specific data for a signaling system. The challenge of these approaches is that it often requires either a computational biology expert or familiarity with sophisticated computational software solutions to extract biological insights from the datasets [47]. A further complication is that genomic or transcriptomic data are often best interpreted in the context of the heterogeneous large-scale datasets that have already been deposited in publicly-available databases [47].

Genomics

Genomic approaches provide the highest number and variety of datasets on human diseases. These approaches include (1) whole-genome or whole-exome sequencing that identify genetic mutations or copy number variations; (2) genome-wide association studies (GWAS) used to identify genetic variants associated with a disease; (3) microarray or RNA-seq techniques for measuring the mRNA or microRNA (miRNA) expression of cells and comparing the levels between states (transcriptomics); and (4) epigenomics analyses focusing on, for example, DNA methylation and its change during differentiation, ageing, and cancer progression. To analyze the genomic datasets of complex diseases, the systems medicine approach is a highly-effective framework to understand the complexity. Disease-related genes may differ among affected individuals, but the affected pathway or network region is likely to be shared [47]. The identified disease-related genes can be used to list potential biomarkers by filtering those specifically relevant to a given disease or disease stage.

In particular, the advent of GWAS identifying candidate susceptibility genes has opened the door to the pathobiology of chronic inflammatory disease. With this, the prospect of a genetic marker for disease diagnosis, prognosis, and therapeutic efficacy in what can otherwise be very heterogeneous diseases is very appealing. GWAS in large populations of patients with chronic inflammatory diseases such as RA can identify common genetic variants that are associated with having that disease [48].

Zhang et al. [49] undertook analysis of the KEGG pathways [50] affected by 11,922 differentially-expressed genes (DEGs), which had been identified by genome-wide association scans in RA patients. The focal adhesion and extracellular matrix receptor interaction pathways were considered high risk RA pathways. Core members of FAC with genetic variants included integrin subunits A and B, actinin, dedicator of cytokinesis 1 (DOCK1), and B cell lymphoma 2 (BCL2). Their data correlate well with the DNA methylome signature in RA, comprising genome-wide DNA methylation loci from fibroblast-like synoviocytes removed at the time of joint replacement from five patients with osteoarthritis and six patients with RA [49]. Nakano et al. [51] undertook global methylation status analysis and identified differential methylation between osteoarthritis and RA in 1206 different genes. Differentially-methylated genes were mapped to KEGG pathways for gene ontology, which highlighted hypomethylation enrichment in the RA sample in loci including genes encoding integrin subunits A and B, actinin, receptor tyrosine kinases, parvin, DOCK1,

and BCL2. Hypomethylation of inflammatory genes has been associated with an increased inflammatory response, as hypomethylation in promoter regions of a gene makes it transcriptionally active [52,53].

Utilizing GWAS-mapped genes or methylome signatures alone for biomarker prediction has its limitations. Firstly, the differential expression of said genes is not assessed. Secondly, the presence or absence of a single polymorphism within a gene may not have a strong enough phenotype to be a useful biomarker [54]. Moreover, the use of methylation status as a biomarker is currently plagued by inaccuracy and poor replication, as there is a need for standardized methods and controls [55].

To overcome the potential limitation of not taking into account differential gene expression, He et al. [56] examined the Gene Expression Omnibus (GEO) microarray data to assess mRNA expression in the specific cell type involved in RA, synovial fibroblasts, to identify DEGs by comparing six RA patients to osteoarthritis patients (an age related, non-autoimmune arthritis) using the linear models for microarray analysis (LIMMA) [57]. The authors undertook functional enrichment of the DEGs using KEGG pathways, with the analysis performed using the database annotation visualization and integrated discovery (DAVID) [58]. Using STRING [59], they created a larger protein–protein interaction (PPI) network for a further functional enrichment, looking for functional complexes using the MCODE plugin for Cytoscape [60]. This multi-layered approach comparing the two types of arthritis identified DEGs for collagen (a predominant member of the extracellular matrix) that were enriched in focal adhesion pathways and extracellular matrix receptor interactions for osteoarthritis, but not RA. The difficulty of biomarker identification based on gene expression studies only is that the studies are often small, thereby not taking into account the rich genetic variability of these complex diseases, and neither gene regulation nor protein levels of DEGs.

Transcriptomics

Combinatorial approaches utilizing DEGs and their regulation have been more successful for biomarker discovery. One mechanism of gene regulation is via small non-coding RNAs (ncRNAs) such as miRNAs. miRNAs function in RNA silencing, by base pairing binding of complementary sequences in mRNAs, thus targeting them for cleavage [61]. In the field of oncology, integrating miRNA, gene expression, and transcription factor signatures has been used to identify biomarkers for papillary thyroid cancer by using pathway enrichment to

identify dysregulated pathways including in focal adhesion [62]. Such approach of integrating miRNA data and differential gene expression for identification of molecular prognostic biomarkers was taken further by Cai and colleagues [63], who identified three potential biomarkers, *CALM2*, miR-19b, and miR181b, for gastric cancer that were related to the FAC and the extracellular matrix receptor. This integrative approach has been, however, less widely used in inflammatory models. For IBD [64] and many other autoimmune diseases including Sjogren's disease [65], we are still at the stage of documenting differential expression levels of miRNAs between disease and control cohorts.

Therefore, despite the central role FAC plays in inflammatory diseases, the number of ncRNAs that could be used as potential biomarkers are still scarce. In the case of UC and Crohn's disease, miRNAs are the most explored ncRNAs in the literature. There is experimental evidence showing elevated levels of specific miRNAs in active UC tissues and in serum [66] .

In recent years, many computational methods emerged that allow the analysis of specific ncRNA–disease associations, predict such connections and select the ones most suitable for experimental validation. For example, heterogeneous graph inference for miRNA–disease association prediction (HGIMDA) [67] and improved random walk with restart for lncRNA-disease association prediction (IRWRLDA) [68] are two viable, novel methods that could be potentially used to describe new targets. HGIMDA constructs a heterogeneous graph out of separate networks: a functional similarity network of miRNAs and a semantic similarity network of diseases, which in combination allowed predicting potential disease–miRNA associations. IRWRLDA uses an improved random walk with restart algorithm on a lncRNA similarity network to rank potentially useful candidate lncRNAs.

Proteomics

Protein biomarker identification is driven by better understanding of the disease processes and signalling pathways involved in perpetuation of pathogenic states. Combining large-scale mass spectrometry (MS)-based proteomics and biological network analysis has been fundamental in the understanding of signalling networks [69], so it stands to reason that using similar techniques may drive biomarker identification for the large datasets that have been proved by proteomic platforms. Like genomics and transcriptomics, biomarker discovery using proteomics has often involved proteome analysis with pathway enrichment. A good example of this is reported by Rukmangadachar and colleagues [70]. They differentiated

intestinal tuberculosis (TB) and ileal Crohn's disease, utilizing MS-based proteome analysis on ileal biopsies of 15 patients, in combination with pathway enrichment using KEGG pathways and the PANTHER annotation resource, and identified biomarkers of both intestinal TB and Crohn's disease. They were able to identify overexpressed proteins in Crohn's disease patients compared to intestinal TB patients. These proteins were annotated to pathways such as the integrin signalling pathway, including a core FAC member, vinculin. However, the proteins they identified were unable to be validated as differential biomarkers in their 52-patient validation cohort using immunohistochemistry. This emphasises the point that a one-step, single-omics approach on a small cohort of patients, whilst identifying potential pathways, lacks the finesse to complete the biomarker discovery.

Systems biology and focal adhesion – the promise for novel biomarker discovery

Looking again at the cancer model, we can see that integrative approaches using both omics data and computational biology have been successful in producing panel biomarkers for cancer subtypes. A good example of this is reported by Zhang and colleagues [71]. They took a systems biology approach to discover, characterize, and validate a panel of breast cancer biomarkers from breast cancer proteomics data. Using liquid chromatography (LC)-coupled MS data from 40 women with breast cancer and 40 women without breast cancer, they identified statistically significant differentially-expressed proteins. They further identified PPI networks and performed pathway analysis with significant literature curation (hypothesis-driven). As a result, they identified a panel of 25 breast cancer biomarkers, which were able to be validated against other proteomic datasets. The top three pathways they identified for the biomarker panel were focal adhesion, regulation of the actin cytoskeleton, as well as complement and coagulation cascades. Combining gene expression data with PPI networks and analysis by a computation network method that utilizes PPI affinity has been equally successful in another breast cancer biomarker discovery study. Protein interactors specific for metastatic breast cancer were identified, which unsurprisingly are part of FAC [72]. Like in cancer, FAC has clearly been implicated in the pathogenesis of complex inflammatory diseases including RA [73] and IBD, leading to the tantalizing possibility of clinical biomarkers identified from within the ranks of FAC.

Utilizing single omics technologies with computation biology has provided potential markers, but these have often failed to stand up to rigorous validation due to small sample

sizes, differences in tissues sampled, or methodological differences. Perhaps a more holistic, integrated approach is needed to meet the needs of modern medicine. This approach towards a more systemic view necessitates obtaining significant insights by adopting a variety of complementary approaches, such as (1) genomics and transcriptional profiling (including miRNA and lncRNA analysis); and (2) functional and phospho-proteomics (affinity purification and MS), as well as other types of large-scale studies, including lipidomics (isolation and MS analysis of lipid content and protein–lipid interactions), chemical proteomics, and compound screening. With the combined and integrated use of these omics approaches, we can identify potential novel biomarkers and drug targets. All biomarkers to be used in clinical practice need independent validation with clinical samples. One such way as used by Szasz et al. [74] is to merge transcriptomic data from multiple independent datasets to cross validate gene expression biomarkers using univariate and multivariate analyses in 1065 patients. Where such samples are not available or not appropriate, clinical trials with patient cohorts need to be undertaken comparing the biomarker candidates identified against a gold standard. An example of this can be seen in Brandse et al. [75] comparing an inflammatory marker, fecal calprotectin, against the gold standard of leukocyte scintigraphy for denoting inflammatory burden in UC.

Conclusions

The FAC is a large, dynamic, multimeric structural and signalling opportunity for biomarker identification. Cancer research has led the way with FAC members being implicated as biomarkers of invasion [76], differentiation between normal and cancer cells [77], prognosis [78], and diagnosis [63]. It is clear that the FAC has a role to play in many inflammatory diseases. However, which member, by which mechanism (be it genomic, transcriptomic, proteomic, or a combinatorial approach with a panel of biomarkers [79]) and in which cell type, remains to be formally validated. Here we presented a few examples of how omics approaches could be exploited, separately or in combination, to provide valuable novel biomarkers for inflammatory diseases from members of the FAC that can undergo further validation in a clinical trial.

Competing interests

The authors certify that they have no conflicts of interest.

Acknowledgments

The work was supported by a Clinical Training Fellowship to JB from the Wellcome Trust, UK and by a fellowship in computational biology to TK at the Earlham Institute, in partnership with the Institute of Food Research, UK, and strategically supported by the Biotechnological and Biosciences Research Council, UK (Grant No. BB/J004529/1).

References

[1] Weaver DR. A1-adenosine receptor gene expression in fetal rat brain. Brain Res Dev Brain Res 1996;94:205–23.

[2] Perlis RH. Translating biomarkers to clinical practice. Mol Psychiatry 2011;16:1076–87.

[3] Pardi DS, Sandborn WJ. Predicting relapse in patients with inflammatory bowel disease: what is the role of biomarkers? Gut 2005;54:321–2.

[4] Geiger B, Spatz JP, Bershadsky AD. Environmental sensing through focal adhesions. Nat Rev Mol Cell Bio 2009;10:21–33.

[5] Jahed Z, Shams H, Mehrbod M, Mofrad MR. Mechanotransduction pathways linking the extracellular matrix to the nucleus. Int Rev Cell Mol Biol 2014;310:171–220.

[6] Schiffhauer ES, Luo T, Mohan K, Srivastava V, Qian X, Griffis ER, et al. Mechanoaccumulative elements of the mammalian actin cytoskeleton. Curr Biol 2016;26:1473–9.

[7] Maruoka M, Sato M, Yuan Y, Ichiba M, Fujii R, Ogawa T, et al. Abl-1-bridged tyrosine phosphorylation of VASP by abelson kinase impairs association of VASP to focal adhesions and regulates leukaemic cell adhesion. Biochem J 2012;441:889–99.

[8] Holle AW, McIntyre AJ, Kehe J, Wijesekara P, Young JL, Vincent LG, et al. High content image analysis of focal adhesion-dependent mechanosensitive stem cell differentiation. Integr Biol (Camb) 2016;8:1049–58.

[9] Case LB, Waterman CM. Integration of actin dynamics and cell adhesion by a three-dimensional, mechanosensitive molecular clutch. Nat Cell Biol 2015;17:955–63.

[10] Goldmann WH. Mechanosensation: a basic cellular process. Prog Mol Biol Transl Sci 2014;126:75–102.

[11] Tachibana M, Kiyokawa E, Hara S, Iemura S, Natsume T, Manabe T, et al. Ankyrin repeat domain 28 (ANKRD28), a novel binding partner of DOCK180, promotes cell migration by regulating focal adhesion formation. Exp Cell Res 2009;315:863–76.

[12] Goicoechea SM, Awadia S, Garcia-Mata R. I'm coming to GEF you: tegulation of RhoGEFs during cell migration. Cell Adh Migr 2014;8:535–49.

[13] Leavitt T, Hu MS, Marshall CD, Barnes LA, Lorenz HP, Longaker MT. Scarless wound healing: finding the right cells and signals. Cell Tissue Res 2016;365:483–93.

[14] You D, Xin J, Volk A, Wei W, Schmidt R, Scurti G, et al. FAK mediates a compensatory survival signal parallel to PI3K-AKT in *PTEN*-null T-ALL cells. Cell Rep 2015;10:2055–68.

[15] Chan PM, Lim L, Manser E. PAK is regulated by PI3K, PIX, CDC42, and PP2Calpha and mediates focal adhesion turnover in the hyperosmotic stress-induced p38 pathway. J Biol Chem 2008;283:24949–61.

[16] Alanko J, Mai A, Jacquemet G, Schauer K, Kaukonen R, Saari M, et al. Integrin endosomal signalling suppresses anoikis. Nat Cell Biol 2015;17:1412–21.

[17] Kato H, Honda S, Yoshida H, Kashiwagi H, Shiraga M, Honma N, et al. SHPS-1 negatively regulates integrin alphaIIbbeta3 function through CD47 without disturbing FAK phosphorylation. J Thromb Haemost 2005;3:763–74.

[18] Nakashima S. Protein kinase C alpha (PKC alpha): regulation and biological function. J Biochem 2002;132:669–75.

[19] Zhao R, Pathak AS, Stouffer GA. Beta(3)-integrin cytoplasmic binding proteins. Arch Immunol Ther Exp (Warsz) 2004;52:348–55.

[20] Abshire MY, Thomas KS, Owen KA, Bouton AH. Macrophage motility requires distinct α5β1/FAK and α4β1/paxillin signaling events. J Leukoc Biol 2011;89:251–7.

[21] Beausejour M, Noel D, Thibodeau S, Bouchard V, Harnois C, Beaulieu JF, et al. Integrin/Fak/Src-mediated regulation of cell survival and anoikis in human intestinal epithelial crypt cells: selective engagement and roles of PI3K isoform complexes. Apoptosis 2012;17:566–78.

[22] Bouchard V, Harnois C, Demers MJ, Thibodeau S, Laquerre V, Gauthier R, et al. B1 integrin/Fak/Src signaling in intestinal epithelial crypt cell survival: integration of complex regulatory mechanisms. Apoptosis 2008;13:531–42.

[23] Zheng Y, Xia Y, Hawke D, Halle M, Tremblay ML, Gao X, et al. FAK phosphorylation by ERK primes ras-induced tyrosine dephosphorylation of FAK mediated by PIN1 and PTP-PEST. Mol Cell 2009;35:11–25.

[24] Zheng Y, Yang W, Xia Y, Hawke D, Liu DX, Lu Z. Ras-induced and extracellular signal-regulated kinase 1 and 2 phosphorylation-dependent isomerization of protein tyrosine phosphatase (PTP)-PEST by PIN1 promotes FAK dephosphorylation by PTP-PEST. Mol Cell Biol 2011;31:4258–69.

[25] Shibanuma M, Mori K, Kim-Kaneyama JR, Nose K. Involvement of FAK and PTP-PEST in the regulation of redox-sensitive nuclear-cytoplasmic shuttling of a LIM protein, Hic-5. Antioxid Redox Signal 2005;7:335–47.

[26] Zaidel-Bar R, Itzkovitz S, Ma'ayan A, Iyengar R, Geiger B. Functional atlas of the integrin adhesome. Nat Cell Biol 2007;9:858–68.

[27] MacGillivray MK, Cruz TF, McCulloch CA. The recruitment of the interleukin-1 (IL-1) receptor-associated kinase (IRAK) into focal adhesion complexes is required for IL-1beta -induced ERK activation. J Biol Chem 2000;275:23509–15.

[28] Wang Q, Wang Y, Downey GP, Plotnikov S, McCulloch CA. A ternary complex comprising FAK, PTPalpha and IP3 receptor 1 functionally engages focal adhesions and the endoplasmic reticulum to mediate IL-1-induced Ca^{2+} signalling in fibroblasts. Biochem J 2016;473:397–410.

[29] Dixit N, Kim MH, Rossaint J, Yamayoshi I, Zarbock A, Simon SI. Leukocyte function antigen-1, kindlin-3, and calcium flux orchestrate neutrophil recruitment during inflammation. J Immunol 2012;189:5954–64.

[30] Klimova Z, Braborec V, Maninova M, Caslavsky J, Weber MJ, Vomastek T. Symmetry breaking in spreading RAT2 fibroblasts requires the MAPK/ERK pathway scaffold RACK1 that integrates FAK, p190A-RhoGAP and ERK2 signaling. Biochim Biophys Acta 2016;1863:2189–200.

[31] Iioka H, Iemura S, Natsume T, Kinoshita N. Wnt signalling regulates paxillin ubiquitination essential for mesodermal cell motility. Nat Cell Biol 2007;9:813–21.

[32] Yu Y, Wu J, Wang Y, Zhao T, Ma B, Liu Y, et al. Kindlin 2 forms a transcriptional complex with beta-catenin and TCF4 to enhance Wnt signalling. EMBO Rep 2012;13:750–8.

[33] Johnson HE, King SJ, Asokan SB, Rotty JD, Bear JE, Haugh JM. F-actin bundles direct the initiation and orientation of lamellipodia through adhesion-based signaling. J Cell Biol 2015;208:443–55.

[34] Fuste NP, Fernandez-Hernandez R, Cemeli T, Mirantes C, Pedraza N, Rafel M, et al. Cytoplasmic cyclin D1 regulates cell invasion and metastasis through the phosphorylation of paxillin. Nat Commun 2016;7:11581.

[35] Shi J, Wu WJ, Hu G, Yu X, Yu GS, Lu H, et al. Regulation of beta-catenin transcription activity by leupaxin in hepatocellular carcinoma. Tumour Biol 2016;37:2313–20.

[36] Salim SY, Soderholm JD. Importance of disrupted intestinal barrier in inflammatory bowel diseases. Inflamm Bowel Dis 2011;17:362–81.

[37] May M, Wang TB, Muller M, Genth H. Difference in F-actin depolymerization induced by toxin B from the *Clostridium difficile* strain VPI 10463 and toxin B from the variant *Clostridium difficile* serotype F strain 1470. Toxins 2013;5:106–19.

[38] Khan MRI, Yazawa T, Anisuzzaman ASM, Semba S, Ma YJ, Uwada J, et al. Activation of focal adhesion kinase via M1 muscarinic acetylcholine receptor is required in restitution of intestinal barrier function after epithelial injury. Bba-Mol Basis Dis 2014;1842:635–45.

[39] Guo SH, Nighot M, Al-Sadi R, Alhmoud T, Nighot P, Ma TY. Lipopolysaccharide regulation of intestinal tight junction permeability is mediated by TLR4 signal transduction pathway activation of FAK and MyD88. J Immunol 2015;195:4999–5010.

[40] Eitel J, Meixenberger K, van Laak C, Orlovski C, Hocke A, Schmeck B, et al. Rac1 regulates the NLRP3 inflammasome which mediates IL-1beta production in *Chlamydophila pneumoniae* infected human mononuclear cells. PLoS One 2012;7:e30379.

[41] Spalinger MR, Kasper S, Gottier C, Lang S, Atrott K, Vavricka SR, et al. NLRP3 tyrosine phosphorylation is controlled by protein tyrosine phosphatase PTPN22. J Clin Invest 2016;126:1783–800.

[42] Nowarski R, Jackson R, Gagliani N, de Zoete MR, Palm NW, Bailis W, et al. Epithelial IL-18 equilibrium controls barrier function in colitis. Cell 2015;163:1444–56.

[43] Thinwa J, Segovia JA, Bose S, Dube PH. Integrin-mediated first signal for inflammasome activation in intestinal epithelial cells. J Immunol 2014;193:1373–82.

[44] Villani AC, Lemire M, Fortin G, Louis E, Silverberg MS, Collette C, et al. Common variants in the *NLRP3* region contribute to Crohn's disease susceptibility. Nat Genet 2009;41:71–6.

[45] Peterson CG, Lampinen M, Hansson T, Liden M, Hallgren R, Carlson M. Evaluation of biomarkers for ulcerative colitis comparing two sampling methods: fecal markers reflect colorectal inflammation both macroscopically and on a cellular level. Scand J Clin Lab Invest 2016:1–9.

[46] Wang E, Zou J, Zaman N, Beitel LK, Trifiro M, Paliouras M. Cancer systems biology in the genome sequencing era: part 1, dissecting and modeling of tumor clones and their networks. Semin Cancer Biol 2013;23:279–85.

[47] Berger B, Peng J, Singh M. Computational solutions for omics data. Nat Rev Genet 2013;14:333–46.

[48] Plenge RM, Criswell LA. Genetic variants that predict response to anti-tumor necrosis factor therapy in rheumatoid arthritis: current challenges and future directions. Curr Opin Rheumatol 2008;20:145–52.

[49] Zhang M, Mu H, Lv H, Duan L, Shang Z, Li J, et al. Integrative analysis of genome-wide association studies and gene expression analysis identifies pathways associated with rheumatoid arthritis. Oncotarget 2016;7:8580–9.

[50] Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000;28:27–30.

[51] Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature in rheumatoid arthritis. Ann Rheum Dis 2013;72:110–7.

[52] Shuto T, Furuta T, Oba M, Xu H, Li JD, Cheung J, et al. Promoter hypomethylation of Toll-like receptor-2 gene is associated with increased proinflammatory response toward bacterial peptidoglycan in cystic fibrosis bronchial epithelial cells. FASEB J 2006;20:782–4.

[53] Barnett M, Bermingham E, McNabb W, Bassett S, Armstrong K, Rounce J, et al. Investigating micronutrients and epigenetic mechanisms in relation to inflammatory bowel disease. Mutat Res 2010;690:71–80.

[54] Chung IM, Ketharnathan S, Thiruvengadam M, Rajakumar G. Rheumatoid arthritis: the stride from research to clinical practice. Int J Mol Sci 2016;17:doi:10.3390/ijms17060900.

[55] Mikeska T, Craig JM. DNA methylation biomarkers: cancer and beyond. Genes (Basel) 2014;5:821–64.

[56] He P, Zhang Z, Liao W, Xu D, Fu M, Kang Y. Screening of gene signatures for rheumatoid arthritis and osteoarthritis based on bioinformatics analysis. Mol Med Rep 2016;14:1587–93.

[57] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.

[58] Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. Genome Biol 2003;4:P3.

[59] von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, et al. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res 2005;33:D433–7.

[60] Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003;4:2.

[61] Ambros V, Lee RC. Identification of microRNAs and other tiny noncoding RNAs by cDNA cloning. Methods Mol Biol 2004;265:131–58.

[62] Ab Mutalib NS, Othman SN, Mohamad Yusof A, Abdullah Suhaimi SN, Muhammad R, Jamal R. Integrated microRNA, gene expression and transcription factors signature in papillary thyroid cancer with lymph node metastasis. PeerJ 2016;4:e2119.

[63] Cai H, Xu J, Han Y, Lu Z, Han T, Ding Y, et al. Integrated miRNA-risk gene-pathway pair network analysis provides prognostic biomarkers for gastric cancer. Onco Targets Ther 2016;9:2975–86.

[64] Xu JW, Li QQ, Tao LL, Cheng YY, Yu J, Chen Q, et al. Involvement of EGFR in the promotion of malignant properties in multidrug resistant breast cancer cells. Int J Oncol 2011;39:1501–9.

[65] Yang Y, Wang JK. The functional analysis of microRNAs involved in NF-kappaB signaling. Eur Rev Med Pharmacol Sci 2016;20:1764–74.

[66] Chapman CG, Pekow J. The emerging role of miRNAs in inflammatory bowel disease: a review. Therap Adv Gastroenterol 2015;8:4–22.

[67] Chen X, Clarence Yan C, Zhang X, You ZH, Huang YA, Yan GY. HGIMDA: heterogeneous graph inference for miRNA-disease association prediction. Oncotarget 2016;7:65257–69.

[68] Chen X, You ZH, Yan GY, Gong DW. IRWRLDA: improved random walk with restart for lncRNA-disease association prediction. Oncotarget 2016;7:57919–31.

[69] Choudhary C, Mann M. Decoding signalling networks by mass spectrometry-based proteomics. Nat Rev Mol Cell Biol 2010;11:427–39.

[70] Rukmangadachar LA, Makharia GK, Mishra A, Das P, Hariprasad G, Srinivasan A, et al. Proteome analysis of the macroscopically affected colonic mucosa of Crohn's disease and intestinal tuberculosis. Sci Rep 2016;6:23162.

[71] Zhang H, Xu Y, Papanastasopoulos P, Stebbing J, Giamas G. Broader implications of SILAC-based proteomics for dissecting signaling dynamics in cancer. Expert Rev Proteomics 2014;11:713–31.

[72] Xin J, Ren X, Chen L, Wang Y. Identifying network biomarkers based on protein–protein interactions and expression data. BMC Med Genomics 2015;8:S11.

[73] Marquez A, Vidal-Bralo L, Rodriguez-Rodriguez L, Gonzalez-Gay MA, Balsa A, Gonzalez-Alvaro I, et al. A combined large-scale meta-analysis identifies *COG6* as a novel shared risk locus for rheumatoid arthritis and systemic lupus erythematosus. Ann Rheum Dis 2017: 76: 286–94.

[74] Szasz AM, Lanczky A, Nagy A, Forster S, Hark K, Green JE, et al. Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. Oncotarget 2016;7:49322–33.

[75] Brandse JF, Bennink RJ, van Eeden S, Lowenberg M, van den Brink GR, D'Haens GR. Performance of common disease activity markers as a reflection of inflammatory burden in ulcerative colitis. Inflamm Bowel Dis 2016;22:1384–90.

[76] Omran OM, Al Sheeha M. Cytoskeletal focal adhesion proteins fascin-1 and paxillin are predictors of malignant progression and poor prognosis in human breast cancer. J Environ Pathol Toxicol Oncol 2015;34:201–12.

[77] Trevino V, Cassese A, Nagy Z, Zhuang X, Herbert J, Antzack P, et al. A network biology approach identifies molecular cross-talk between normal prostate epithelial and prostate carcinoma cells. PLoS Comput Biol 2016;12:e1004884.

[78] He Y, Shao F, Pi W, Shi C, Chen Y, Gong D, et al. Largescale transcriptomics analysis suggests over-expression of BGH3, MMP9 and PDIA3 in oral squamous cell carcinoma. PLoS One 2016;11:e0146530.

[79] Plevy S, Silverberg MS, Lockton S, Stockfisch T, Croner L, Stachelski J, et al. Combined serological, genetic, and inflammatory markers differentiate non-IBD, Crohn's disease, and ulcerative colitis patients. Inflamm Bowel Dis 2013;19:1139–48.

Figure legends

Figure 1 Potential sites of biomarker used in IBD

IBD-U refers to IBD-undefined, for which the patient's endoscopy and histology cannot give a clear distinction between UC and CD. IBD, inflammatory bowel disease; UC, ulcerative colitis; CD, Crohn's disease.

Figure 2 Omics approaches with complementary potential to be integrated

Genomics, transcriptomics, and proteomics approaches can be used to identify and discover the detailed component, mechanisms, and regulation of the FAC members in normal and in diseased states. The differential analysis is capable to point out novel biomarkers. FAC, focal adhesion complex; GWAS, genome-wide association studies; MS, mass spectrometry; WES, whole-exome sequencing; WGS, whole-genome sequencing.

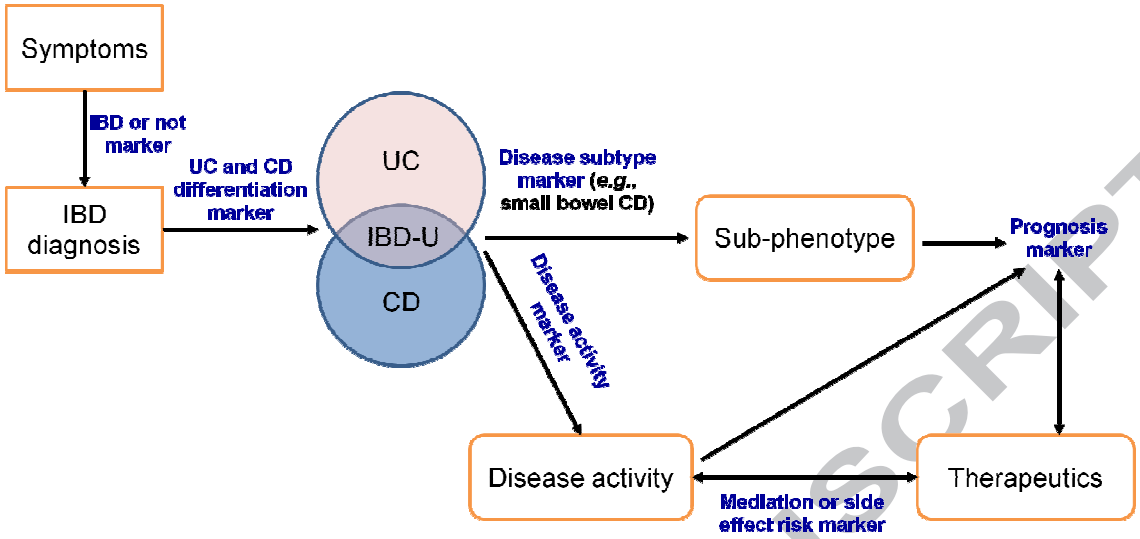
Figure 3 Biomarker related publications about focal adhesion complex

We compared the total publications in PubMed identified with MeSH terms 'focal adhesion and biomarker' with 'cancer, focal adhesion and biomarker' from the last 20 years. The figure highlights the unchanged and low number of biomarker-related studies involving the FAC and non-cancer diseases compared to cancer related studies. FAC, focal adhesion complex.

Tables

Table 1 Component examples of the focal adhesion complex

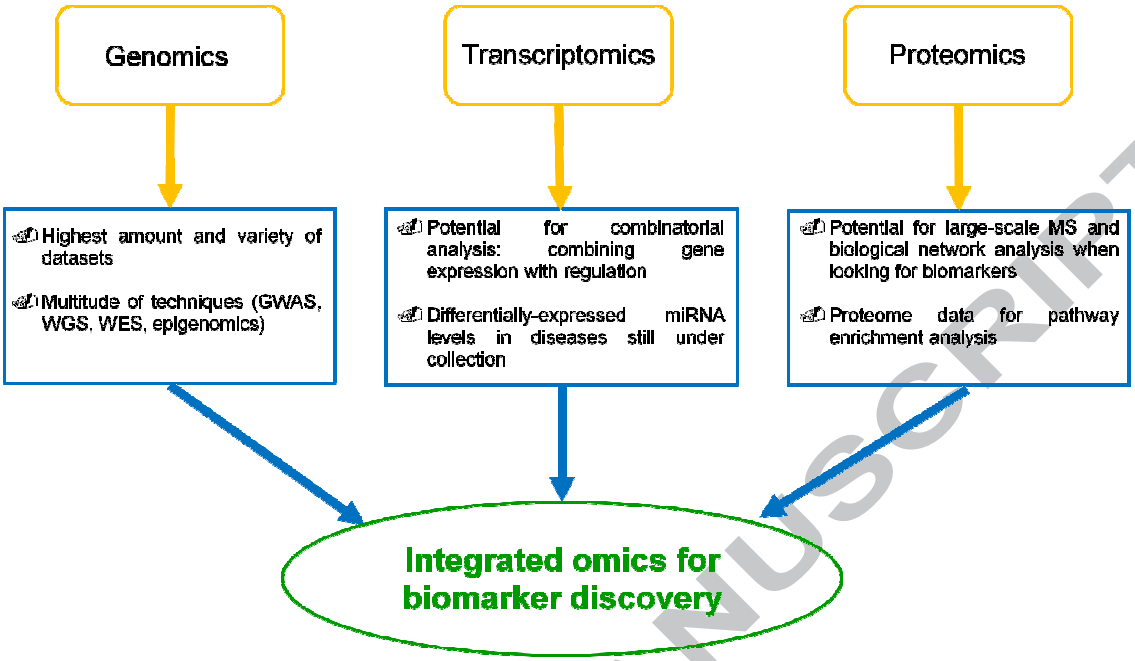
554



555

556

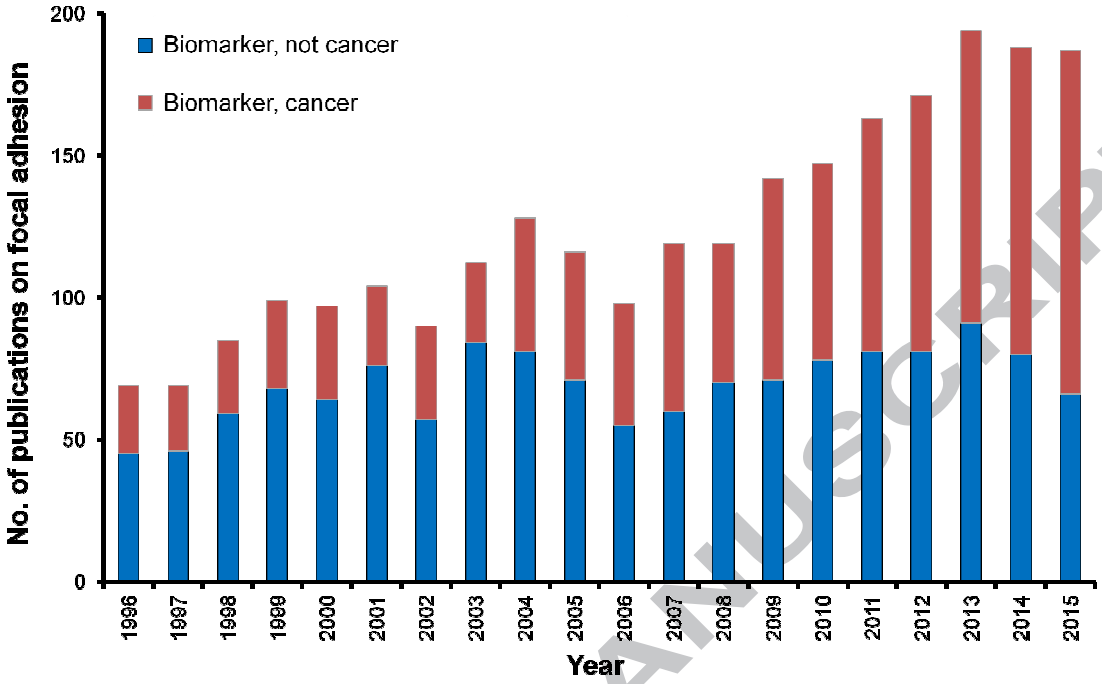
557



558

559

560



561

562

563 **Table 1 Component examples of the focal adhesion complex**

Category	Example	Function	Refs.
Actin binding	Actinin1, filamin A, cortactin, zyxin	Crosslink actin; remodel cytoskeleton	[5,6]
Adapter	SORBS1, ABI1	Link proximal signal pathways; facilitate signal transduction	[7,8]
Cytoskeletal	Actin, vinculin, plectin, ezrin, paxillin	Facilitate and stabilize signalling platforms; remodel cell shape and movement	[9,10]
GAP/GEF	DOCK1, ELMO1	Activate small GTPases	[11]
GTPase	Rac1, RhoA	Signal cytoskeletal remodelling, cell growth, phagocytosis, and ruffled borders	[12]
Metalloproteinase	ADAM12	Disintegrin	[13]
PIK/phosphatase	PI3K, INPPL1, PTEN	Regulate AKT/PKB signalling pathway; regulate signalling via IRS proteins	[14,15]
Receptor	Integrins, IL1R, CD47	Bind to ligands for extracellular matrix constituents including fibronectin and thrombospondin	[16,17]
Serine/threonine kinase	PAK1, AKT, PRKCA	Effectors linking Rho GTPases to cytoskeletal reorganization; phosphorylate BCL2	[15,18]
Transcription factor	ITGB3BP	Transcriptional co-regulator	[19]
Tyrosine kinase	FAK, SYK, SRC	Regulate FAC assembly and disassembly	[20–22]
Tyrosine phosphatase	PTPN1, 2, 6, 11, 12, 22, PTP-PEST	Regulate maturation of focal adhesion; recruit signalling molecules	[23–25]

564 *Note*: GTPase, guanosine triphosphatase; GAP, GTPase activating protein; GEF,
565 **guanine nucleotide exchange** factor; PIK, phosphoinositide kinase; SORBS1, sorbin and
566 SH3 domain-containing 1; ABI1, Abelson interactor 1; DOCK1, dedicator of cytokinesis
567 protein 1; ELMO1, engulfment and cell motility protein 1; Rac1, Ras-related C3 botulinum
568 toxin substrate 1; RhoA, Ras homolog gene family, member A; ADAM12, disintegrin and
569 metalloproteinase domain-containing protein 12; PI3K, phosphatidyl-inositol-3-kinase;
570 INPPL1, inositol polyphosphate phosphatase like 1; PTEN, phosphatase and tensin homolog;
571 IL1R, interleukin-1 receptor type 1; CD47, Cluster of Differentiation 47; PAK1,
572 serine/threonine-protein kinase 1; AKT, RAC-alpha serine/threonine-protein kinase; PRKCA,
573 protein kinase C alpha type; ITGB3BP, integrin subunit beta 3 binding protein; FAC, focal
574 adhesion complex; FAK, focal adhesion kinase; SYK, spleen tyrosine kinase; SRC, proto-
575 oncogene tyrosine-protein kinase; PTPN, tyrosine-protein phosphatase non-receptor type;
576 PTP, protein tyrosine phosphatase; PKB, protein kinase B; Bcl2, B-cell lymphoma 2.
577