

## Accepted Manuscript

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

Johanne Brooks, Alastair Watson, Tamas Korcsmaros

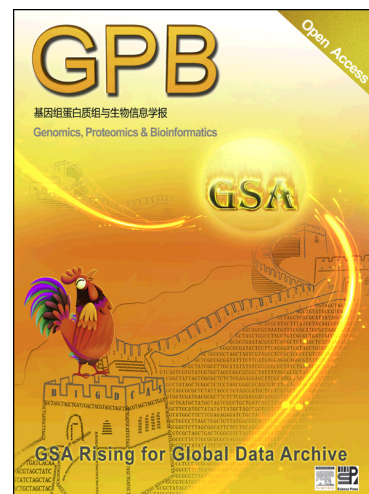
PII: S1672-0229(17)30041-4  
DOI: <http://dx.doi.org/10.1016/j.gpb.2016.12.003>  
Reference: GPB 240

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>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



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

3

4 Johanne Brooks<sup>1,2,3,a</sup>, Alastair Watson<sup>1,2,3,b</sup>, Tamas Korcsmaros<sup>1,4,\*c</sup>

5

6 <sup>1</sup> *Gut Health and Food Safety Institute Strategic Programme, Institute of Food Research,*  
7 *Norwich Research Park, Norwich NR4 7UA, United Kingdom*

8 <sup>2</sup> *Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich NR4*  
9 *7TJ, United Kingdom*

10 <sup>3</sup> *Gastroenterology Department, Norfolk and Norwich University Hospital, Norwich NR4*  
11 *7UY, United Kingdom*

12 <sup>4</sup> *Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, United Kingdom*

13

14 \*Corresponding author.

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

16

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

18

19 <sup>a</sup> ORCID: 0000-0001-5967-7225.

20 <sup>b</sup> ORCID: 0000-0003-3326-0426.

21 <sup>c</sup> ORCID: 0000-0003-1717-996X.

22

23 Total count of words: 5928

24 Total count of figures: 3

25 Total count of tables: 1

26

27 **Abstract**

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

41

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

## 43 Introduction

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

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

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

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

78 The focal adhesion function is both mechanical and responsive. It is mechanical in terms  
79 of anchoring the cell to the extracellular matrix via binding of integrins to their extracellular  
80 ligands and to the actin cytoskeleton to modify the physical and topographical characteristics  
81 of the cell. This has direct implications for wound healing as well as invasion and the  
82 metastatic nature of the cancer cell. The responsive function of the FAC is diverse and multi-  
83 layered. Depending on the initiating signal, FAC can be involved in regulating inflammatory  
84 gene expression via signal transduction pathways such as interleukin 1 (IL-1) signalling  
85 [27,28] or regulating calcium fluxes via phosphatidyl inositol signalling [29], which impact  
86 on inflammatory cascades. Many molecules in the FAC are involved in downstream  
87 signalling pathways, for instance, the MAPK/ERK pathway [30], AKT1 [22], and Wnt  
88 signalling [31,32]. In this way, pathways impacted by the FAC are as varied as apoptosis  
89 [21], production of cellular protrusions [33], cell cycle progression [34], and cell proliferation  
90 [35].

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

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

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

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

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

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

144

145 **Genomics**

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

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

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



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

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

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

202

### 203 **Transcriptomics**

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



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

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

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

233

## 234 **Proteomics**

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

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

253

### 254 **Systems biology and focal adhesion – the promise for novel biomarker** 255 **discovery**

256 Looking again at the cancer model, we can see that integrative approaches using both omics  
257 data and computational biology have been successful in producing panel biomarkers for  
258 cancer subtypes. A good example of this is reported by Zhang and colleagues [71]. They took  
259 a systems biology approach to discover, characterize, and validate a panel of breast cancer  
260 biomarkers from breast cancer proteomics data. Using liquid chromatography (LC)-coupled  
261 MS data from 40 women with breast cancer and 40 women without breast cancer, they  
262 identified statistically significant differentially-expressed proteins. They further identified  
263 PPI networks and performed pathway analysis with significant literature curation  
264 (hypothesis-driven). As a result, they identified a panel of 25 breast cancer biomarkers, which  
265 were able to be validated against other proteomic datasets. The top three pathways they  
266 identified for the biomarker panel were focal adhesion, regulation of the actin cytoskeleton,  
267 as well as complement and coagulation cascades. Combining gene expression data with PPI  
268 networks and analysis by a computation network method that utilizes PPI affinity has been  
269 equally successful in another breast cancer biomarker discovery study. Protein interactors  
270 specific for metastatic breast cancer were identified, which unsurprisingly are part of FAC  
271 [72]. Like in cancer, FAC has clearly been implicated in the pathogenesis of complex  
272 inflammatory diseases including RA [73] and IBD, leading to the tantalizing possibility of  
273 clinical biomarkers identified from within the ranks of FAC.

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

276 sizes, differences in tissues sampled, or methodological differences. Perhaps a more holistic,  
277 integrated approach is needed to meet the needs of modern medicine. This approach towards  
278 a more systemic view necessitates obtaining significant insights by adopting a variety of  
279 complementary approaches, such as (1) genomics and transcriptional profiling (including  
280 miRNA and lncRNA analysis); and (2) functional and phospho-proteomics (affinity  
281 purification and MS), as well as other types of large-scale studies, including lipidomics  
282 (isolation and MS analysis of lipid content and protein–lipid interactions), chemical  
283 proteomics, and compound screening. With the combined and integrated use of these omics  
284 approaches, we can identify potential novel biomarkers and drug targets. All biomarkers to be  
285 used in clinical practice need independent validation with clinical samples. One such way as  
286 used by Szasz et al. [74] is to merge transcriptomic data from multiple independent datasets  
287 to cross validate gene expression biomarkers using univariate and multivariate analyses in  
288 1065 patients. Where such samples are not available or not appropriate, clinical trials with  
289 patient cohorts need to be undertaken comparing the biomarker candidates identified against  
290 a gold standard. An example of this can be seen in Brandse et al. [75] comparing an  
291 inflammatory marker, fecal calprotectin, against the gold standard of leukocyte scintigraphy  
292 for denoting inflammatory burden in UC.

293

## 294 **Conclusions**

295 The FAC is a large, dynamic, multimeric structural and signalling opportunity for biomarker  
296 identification. Cancer research has led the way with FAC members being implicated as  
297 biomarkers of invasion [76], differentiation between normal and cancer cells [77], prognosis  
298 [78], and diagnosis [63]. It is clear that the FAC has a role to play in many inflammatory  
299 diseases. However, which member, by which mechanism (be it genomic, transcriptomic,  
300 proteomic, or a combinatorial approach with a panel of biomarkers [79]) and in which cell  
301 type, remains to be formally validated. Here we presented a few examples of how omics  
302 approaches could be exploited, separately or in combination, to provide valuable novel  
303 biomarkers for inflammatory diseases from members of the FAC that can undergo further  
304 validation in a clinical trial.

305

## 306 **Competing interests**

307 The authors certify that they have no conflicts of interest.

308

309 **Acknowledgments**

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

314

315 **References**

- 316 [1] Weaver DR. A1-adenosine receptor gene expression in fetal rat brain. *Brain Res Dev*  
317 *Brain Res* 1996;94:205–23.
- 318 [2] Perlis RH. Translating biomarkers to clinical practice. *Mol Psychiatry* 2011;16:1076–87.
- 319 [3] Pardi DS, Sandborn WJ. Predicting relapse in patients with inflammatory bowel disease:  
320 what is the role of biomarkers? *Gut* 2005;54:321–2.
- 321 [4] Geiger B, Spatz JP, Bershadsky AD. Environmental sensing through focal adhesions. *Nat*  
322 *Rev Mol Cell Bio* 2009;10:21–33.
- 323 [5] Jahed Z, Shams H, Mehrbod M, Mofrad MR. Mechanotransduction pathways linking the  
324 extracellular matrix to the nucleus. *Int Rev Cell Mol Biol* 2014;310:171–220.
- 325 [6] Schiffhauer ES, Luo T, Mohan K, Srivastava V, Qian X, Griffis ER, et al.  
326 Mechanoaccumulative elements of the mammalian actin cytoskeleton. *Curr Biol*  
327 2016;26:1473–9.
- 328 [7] Maruoka M, Sato M, Yuan Y, Ichiba M, Fujii R, Ogawa T, et al. Abl-1-bridged tyrosine  
329 phosphorylation of VASP by abelson kinase impairs association of VASP to focal  
330 adhesions and regulates leukaemic cell adhesion. *Biochem J* 2012;441:889–99.
- 331 [8] Holle AW, McIntyre AJ, Kehe J, Wijesekara P, Young JL, Vincent LG, et al. High  
332 content image analysis of focal adhesion-dependent mechanosensitive stem cell  
333 differentiation. *Integr Biol (Camb)* 2016;8:1049–58.
- 334 [9] Case LB, Waterman CM. Integration of actin dynamics and cell adhesion by a three-  
335 dimensional, mechanosensitive molecular clutch. *Nat Cell Biol* 2015;17:955–63.
- 336 [10] Goldmann WH. Mechanosensation: a basic cellular process. *Prog Mol Biol Transl Sci*  
337 2014;126:75–102.
- 338 [11] Tachibana M, Kiyokawa E, Hara S, Iemura S, Natsume T, Manabe T, et al. Ankyrin  
339 repeat domain 28 (ANKRD28), a novel binding partner of DOCK180, promotes cell  
340 migration by regulating focal adhesion formation. *Exp Cell Res* 2009;315:863–76.
- 341 [12] Goicoechea SM, Awadia S, Garcia-Mata R. I'm coming to GEF you: regulation of  
342 RhoGEFs during cell migration. *Cell Adh Migr* 2014;8:535–49.
- 343 [13] Leavitt T, Hu MS, Marshall CD, Barnes LA, Lorenz HP, Longaker MT. Scarless wound  
344 healing: finding the right cells and signals. *Cell Tissue Res* 2016;365:483–93.
- 345 [14] You D, Xin J, Volk A, Wei W, Schmidt R, Scurti G, et al. FAK mediates a compensatory  
346 survival signal parallel to PI3K-AKT in *PTEN*-null T-ALL cells. *Cell Rep*  
347 2015;10:2055–68.

- 348 [15]Chan PM, Lim L, Manser E. PAK is regulated by PI3K, PIX, CDC42, and PP2Calpha  
349 and mediates focal adhesion turnover in the hyperosmotic stress-induced p38 pathway. *J*  
350 *Biol Chem* 2008;283:24949–61.
- 351 [16]Alanko J, Mai A, Jacquemet G, Schauer K, Kaukonen R, Saari M, et al. Integrin  
352 endosomal signalling suppresses anoikis. *Nat Cell Biol* 2015;17:1412–21.
- 353 [17]Kato H, Honda S, Yoshida H, Kashiwagi H, Shiraga M, Honma N, et al. SHPS-1  
354 negatively regulates integrin  $\alpha$ IIb $\beta$ 3 function through CD47 without disturbing  
355 FAK phosphorylation. *J Thromb Haemost* 2005;3:763–74.
- 356 [18]Nakashima S. Protein kinase C alpha (PKC alpha): regulation and biological function. *J*  
357 *Biochem* 2002;132:669–75.
- 358 [19]Zhao R, Pathak AS, Stouffer GA. Beta(3)-integrin cytoplasmic binding proteins. *Arch*  
359 *Immunol Ther Exp (Warsz)* 2004;52:348–55.
- 360 [20]Abshire MY, Thomas KS, Owen KA, Bouton AH. Macrophage motility requires distinct  
361  $\alpha$ 5 $\beta$ 1/FAK and  $\alpha$ 4 $\beta$ 1/paxillin signaling events. *J Leukoc Biol* 2011;89:251–7.
- 362 [21]Beausejour M, Noel D, Thibodeau S, Bouchard V, Harnois C, Beaulieu JF, et al.  
363 Integrin/Fak/Src-mediated regulation of cell survival and anoikis in human intestinal  
364 epithelial crypt cells: selective engagement and roles of PI3K isoform complexes.  
365 *Apoptosis* 2012;17:566–78.
- 366 [22]Bouchard V, Harnois C, Demers MJ, Thibodeau S, Laquerre V, Gauthier R, et al. B1  
367 integrin/Fak/Src signaling in intestinal epithelial crypt cell survival: integration of  
368 complex regulatory mechanisms. *Apoptosis* 2008;13:531–42.
- 369 [23]Zheng Y, Xia Y, Hawke D, Halle M, Tremblay ML, Gao X, et al. FAK phosphorylation  
370 by ERK primes ras-induced tyrosine dephosphorylation of FAK mediated by PIN1 and  
371 PTP-PEST. *Mol Cell* 2009;35:11–25.
- 372 [24]Zheng Y, Yang W, Xia Y, Hawke D, Liu DX, Lu Z. Ras-induced and extracellular  
373 signal-regulated kinase 1 and 2 phosphorylation-dependent isomerization of protein  
374 tyrosine phosphatase (PTP)-PEST by PIN1 promotes FAK dephosphorylation by PTP-  
375 PEST. *Mol Cell Biol* 2011;31:4258–69.
- 376 [25]Shibanuma M, Mori K, Kim-Kaneyama JR, Nose K. Involvement of FAK and PTP-  
377 PEST in the regulation of redox-sensitive nuclear-cytoplasmic shuttling of a LIM  
378 protein, Hic-5. *Antioxid Redox Signal* 2005;7:335–47.
- 379 [26]Zaidel-Bar R, Itzkovitz S, Ma'ayan A, Iyengar R, Geiger B. Functional atlas of the  
380 integrin adhesome. *Nat Cell Biol* 2007;9:858–68.
- 381 [27]MacGillivray MK, Cruz TF, McCulloch CA. The recruitment of the interleukin-1 (IL-1)  
382 receptor-associated kinase (IRAK) into focal adhesion complexes is required for IL-  
383 1beta -induced ERK activation. *J Biol Chem* 2000;275:23509–15.
- 384 [28]Wang Q, Wang Y, Downey GP, Plotnikov S, McCulloch CA. A ternary complex  
385 comprising FAK, PTPalpha and IP3 receptor 1 functionally engages focal adhesions and  
386 the endoplasmic reticulum to mediate IL-1-induced  $Ca^{2+}$  signalling in fibroblasts.  
387 *Biochem J* 2016;473:397–410.
- 388 [29]Dixit N, Kim MH, Rossaint J, Yamayoshi I, Zarbock A, Simon SI. Leukocyte function  
389 antigen-1, kindlin-3, and calcium flux orchestrate neutrophil recruitment during  
390 inflammation. *J Immunol* 2012;189:5954–64.

- 391 [30]Klimova Z, Braborec V, Maninova M, Caslavsky J, Weber MJ, Vomastek T. Symmetry  
392 breaking in spreading RAT2 fibroblasts requires the MAPK/ERK pathway scaffold  
393 RACK1 that integrates FAK, p190A-RhoGAP and ERK2 signaling. *Biochim Biophys*  
394 *Acta* 2016;1863:2189–200.
- 395 [31]Iioka H, Iemura S, Natsume T, Kinoshita N. Wnt signalling regulates paxillin  
396 ubiquitination essential for mesodermal cell motility. *Nat Cell Biol* 2007;9:813–21.
- 397 [32]Yu Y, Wu J, Wang Y, Zhao T, Ma B, Liu Y, et al. Kindlin 2 forms a transcriptional  
398 complex with beta-catenin and TCF4 to enhance Wnt signalling. *EMBO Rep*  
399 2012;13:750–8.
- 400 [33]Johnson HE, King SJ, Asokan SB, Rotty JD, Bear JE, Haugh JM. F-actin bundles direct  
401 the initiation and orientation of lamellipodia through adhesion-based signaling. *J Cell*  
402 *Biol* 2015;208:443–55.
- 403 [34]Fuste NP, Fernandez-Hernandez R, Cemeli T, Mirantes C, Pedraza N, Rafel M, et al.  
404 Cytoplasmic cyclin D1 regulates cell invasion and metastasis through the  
405 phosphorylation of paxillin. *Nat Commun* 2016;7:11581.
- 406 [35]Shi J, Wu WJ, Hu G, Yu X, Yu GS, Lu H, et al. Regulation of beta-catenin transcription  
407 activity by leupaxin in hepatocellular carcinoma. *Tumour Biol* 2016;37:2313–20.
- 408 [36]Salim SY, Soderholm JD. Importance of disrupted intestinal barrier in inflammatory  
409 bowel diseases. *Inflamm Bowel Dis* 2011;17:362–81.
- 410 [37]May M, Wang TB, Muller M, Genth H. Difference in F-actin depolymerization induced  
411 by toxin B from the *Clostridium difficile* strain VPI 10463 and toxin B from the variant  
412 *Clostridium difficile* serotype F strain 1470. *Toxins* 2013;5:106–19.
- 413 [38]Khan MRI, Yazawa T, Anisuzzaman ASM, Semba S, Ma YJ, Uwada J, et al. Activation  
414 of focal adhesion kinase via M1 muscarinic acetylcholine receptor is required in  
415 restitution of intestinal barrier function after epithelial injury. *Bba-Mol Basis Dis*  
416 2014;1842:635–45.
- 417 [39]Guo SH, Nighot M, Al-Sadi R, Alhmoud T, Nighot P, Ma TY. Lipopolysaccharide  
418 regulation of intestinal tight junction permeability is mediated by TLR4 signal  
419 transduction pathway activation of FAK and MyD88. *J Immunol* 2015;195:4999–5010.
- 420 [40]Eitel J, Meixenberger K, van Laak C, Orlovski C, Hocke A, Schmeck B, et al. Rac1  
421 regulates the NLRP3 inflammasome which mediates IL-1beta production in  
422 *Chlamydomphila pneumoniae* infected human mononuclear cells. *PLoS One*  
423 2012;7:e30379.
- 424 [41]Spalinger MR, Kasper S, Gottier C, Lang S, Atrott K, Vavricka SR, et al. NLRP3  
425 tyrosine phosphorylation is controlled by protein tyrosine phosphatase PTPN22. *J Clin*  
426 *Invest* 2016;126:1783–800.
- 427 [42]Nowarski R, Jackson R, Gagliani N, de Zoete MR, Palm NW, Bailis W, et al. Epithelial  
428 IL-18 equilibrium controls barrier function in colitis. *Cell* 2015;163:1444–56.
- 429 [43]Thinwa J, Segovia JA, Bose S, Dube PH. Integrin-mediated first signal for  
430 inflammasome activation in intestinal epithelial cells. *J Immunol* 2014;193:1373–82.
- 431 [44]Villani AC, Lemire M, Fortin G, Louis E, Silverberg MS, Collette C, et al. Common  
432 variants in the *NLRP3* region contribute to Crohn's disease susceptibility. *Nat Genet*  
433 2009;41:71–6.



- 434 [45]Peterson CG, Lampinen M, Hansson T, Liden M, Hallgren R, Carlson M. Evaluation of  
435 biomarkers for ulcerative colitis comparing two sampling methods: fecal markers reflect  
436 colorectal inflammation both macroscopically and on a cellular level. *Scand J Clin Lab*  
437 *Invest* 2016;1–9.
- 438 [46]Wang E, Zou J, Zaman N, Beitel LK, Trifiro M, Paliouras M. Cancer systems biology in  
439 the genome sequencing era: part 1, dissecting and modeling of tumor clones and their  
440 networks. *Semin Cancer Biol* 2013;23:279–85.
- 441 [47]Berger B, Peng J, Singh M. Computational solutions for omics data. *Nat Rev Genet*  
442 2013;14:333–46.
- 443 [48]Plenge RM, Criswell LA. Genetic variants that predict response to anti-tumor necrosis  
444 factor therapy in rheumatoid arthritis: current challenges and future directions. *Curr Opin*  
445 *Rheumatol* 2008;20:145–52.
- 446 [49]Zhang M, Mu H, Lv H, Duan L, Shang Z, Li J, et al. Integrative analysis of genome-wide  
447 association studies and gene expression analysis identifies pathways associated with  
448 rheumatoid arthritis. *Oncotarget* 2016;7:8580–9.
- 449 [50]Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic*  
450 *Acids Res* 2000;28:27–30.
- 451 [51]Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature  
452 in rheumatoid arthritis. *Ann Rheum Dis* 2013;72:110–7.
- 453 [52]Shuto T, Furuta T, Oba M, Xu H, Li JD, Cheung J, et al. Promoter hypomethylation of  
454 Toll-like receptor-2 gene is associated with increased proinflammatory response toward  
455 bacterial peptidoglycan in cystic fibrosis bronchial epithelial cells. *FASEB J*  
456 2006;20:782–4.
- 457 [53]Barnett M, Bermingham E, McNabb W, Bassett S, Armstrong K, Rounce J, et al.  
458 Investigating micronutrients and epigenetic mechanisms in relation to inflammatory  
459 bowel disease. *Mutat Res* 2010;690:71–80.
- 460 [54]Chung IM, Ketharnathan S, Thiruvengadam M, Rajakumar G. Rheumatoid arthritis: the  
461 stride from research to clinical practice. *Int J Mol Sci*  
462 2016;17:doi:10.3390/ijms17060900.
- 463 [55]Mikeska T, Craig JM. DNA methylation biomarkers: cancer and beyond. *Genes (Basel)*  
464 2014;5:821–64.
- 465 [56]He P, Zhang Z, Liao W, Xu D, Fu M, Kang Y. Screening of gene signatures for  
466 rheumatoid arthritis and osteoarthritis based on bioinformatics analysis. *Mol Med Rep*  
467 2016;14:1587–93.
- 468 [57]Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential  
469 expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*  
470 2015;43:e47.
- 471 [58]Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID:  
472 database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4:P3.
- 473 [59]von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, et al. STRING:  
474 known and predicted protein-protein associations, integrated and transferred across  
475 organisms. *Nucleic Acids Res* 2005;33:D433–7.
- 476 [60]Bader GD, Hogue CW. An automated method for finding molecular complexes in large  
477 protein interaction networks. *BMC Bioinformatics* 2003;4:2.



- 478 [61]Ambros V, Lee RC. Identification of microRNAs and other tiny noncoding RNAs by  
479 cDNA cloning. *Methods Mol Biol* 2004;265:131–58.
- 480 [62]Ab Mutalib NS, Othman SN, Mohamad Yusof A, Abdullah Suhaimi SN, Muhammad R,  
481 Jamal R. Integrated microRNA, gene expression and transcription factors signature in  
482 papillary thyroid cancer with lymph node metastasis. *PeerJ* 2016;4:e2119.
- 483 [63]Cai H, Xu J, Han Y, Lu Z, Han T, Ding Y, et al. Integrated miRNA-risk gene-pathway  
484 pair network analysis provides prognostic biomarkers for gastric cancer. *Onco Targets*  
485 *Ther* 2016;9:2975–86.
- 486 [64]Xu JW, Li QQ, Tao LL, Cheng YY, Yu J, Chen Q, et al. Involvement of EGFR in the  
487 promotion of malignant properties in multidrug resistant breast cancer cells. *Int J Oncol*  
488 2011;39:1501–9.
- 489 [65]Yang Y, Wang JK. The functional analysis of microRNAs involved in NF-kappaB  
490 signaling. *Eur Rev Med Pharmacol Sci* 2016;20:1764–74.
- 491 [66]Chapman CG, Pekow J. The emerging role of miRNAs in inflammatory bowel disease: a  
492 review. *Therap Adv Gastroenterol* 2015;8:4–22.
- 493 [67]Chen X, Clarence Yan C, Zhang X, You ZH, Huang YA, Yan GY. HGIMDA:  
494 heterogeneous graph inference for miRNA-disease association prediction. *Oncotarget*  
495 2016;7:65257–69.
- 496 [68]Chen X, You ZH, Yan GY, Gong DW. IRWRLDA: improved random walk with restart  
497 for lncRNA-disease association prediction. *Oncotarget* 2016;7:57919–31.
- 498 [69]Choudhary C, Mann M. Decoding signalling networks by mass spectrometry-based  
499 proteomics. *Nat Rev Mol Cell Biol* 2010;11:427–39.
- 500 [70]Rukmangadachar LA, Makharia GK, Mishra A, Das P, Hariprasad G, Srinivasan A, et al.  
501 Proteome analysis of the macroscopically affected colonic mucosa of Crohn's disease  
502 and intestinal tuberculosis. *Sci Rep* 2016;6:23162.
- 503 [71]Zhang H, Xu Y, Papanastasopoulos P, Stebbing J, Giamas G. Broader implications of  
504 SILAC-based proteomics for dissecting signaling dynamics in cancer. *Expert Rev*  
505 *Proteomics* 2014;11:713–31.
- 506 [72]Xin J, Ren X, Chen L, Wang Y. Identifying network biomarkers based on protein-  
507 protein interactions and expression data. *BMC Med Genomics* 2015;8:S11.
- 508 [73]Marquez A, Vidal-Bralo L, Rodriguez-Rodriguez L, Gonzalez-Gay MA, Balsa A,  
509 Gonzalez-Alvaro I, et al. A combined large-scale meta-analysis identifies *COG6* as a  
510 novel shared risk locus for rheumatoid arthritis and systemic lupus erythematosus. *Ann*  
511 *Rheum Dis* 2017; 76: 286–94.
- 512 [74]Szasz AM, Lanczky A, Nagy A, Forster S, Hark K, Green JE, et al. Cross-validation of  
513 survival associated biomarkers in gastric cancer using transcriptomic data of 1,065  
514 patients. *Oncotarget* 2016;7:49322–33.
- 515 [75]Brandse JF, Bennink RJ, van Eeden S, Lowenberg M, van den Brink GR, D'Haens GR.  
516 Performance of common disease activity markers as a reflection of inflammatory burden  
517 in ulcerative colitis. *Inflamm Bowel Dis* 2016;22:1384–90.
- 518 [76]Omran OM, Al Sheeha M. Cytoskeletal focal adhesion proteins fascin-1 and paxillin are  
519 predictors of malignant progression and poor prognosis in human breast cancer. *J*  
520 *Environ Pathol Toxicol Oncol* 2015;34:201–12.

521 [77]Trevino V, Cassese A, Nagy Z, Zhuang X, Herbert J, Antzack P, et al. A network  
522 biology approach identifies molecular cross-talk between normal prostate epithelial and  
523 prostate carcinoma cells. PLoS Comput Biol 2016;12:e1004884.

524 [78]He Y, Shao F, Pi W, Shi C, Chen Y, Gong D, et al. Largescale transcriptomics analysis  
525 suggests over-expression of BGH3, MMP9 and PDIA3 in oral squamous cell carcinoma.  
526 PLoS One 2016;11:e0146530.

527 [79]Plevy S, Silverberg MS, Lockton S, Stockfisch T, Croner L, Stachelski J, et al.  
528 Combined serological, genetic, and inflammatory markers differentiate non-IBD,  
529 Crohn's disease, and ulcerative colitis patients. Inflamm Bowel Dis 2013;19:1139–48.

530

### 531 **Figure legends**

#### 532 **Figure 1 Potential sites of biomarker used in IBD**

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

536

#### 537 **Figure 2 Omics approaches with complementary potential to be integrated**

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

543

#### 544 **Figure 3 Biomarker related publications about focal adhesion complex**

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

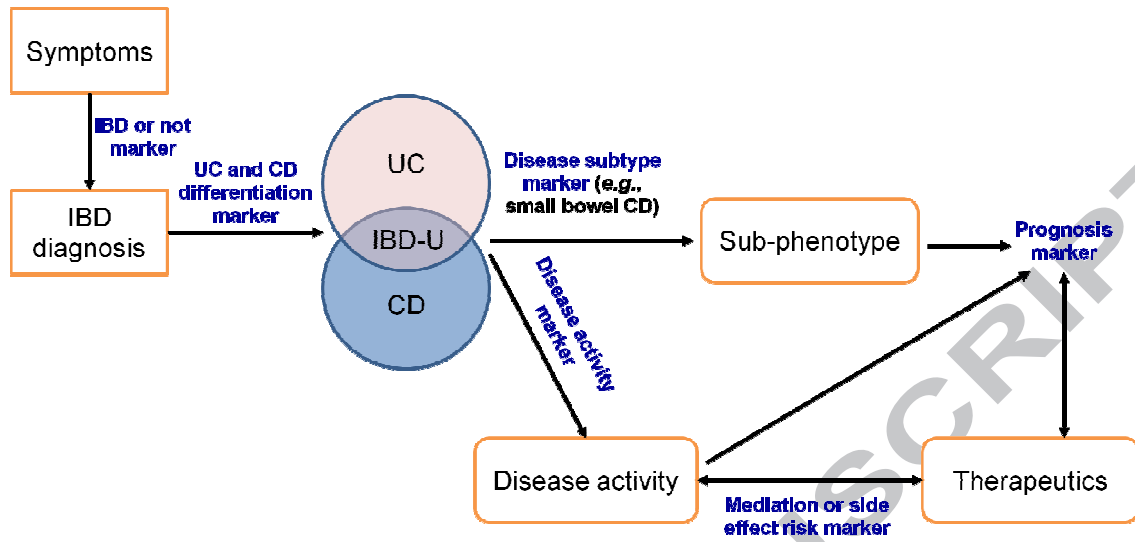
550

### 551 **Tables**

#### 552 **Table 1 Component examples of the focal adhesion complex**

553

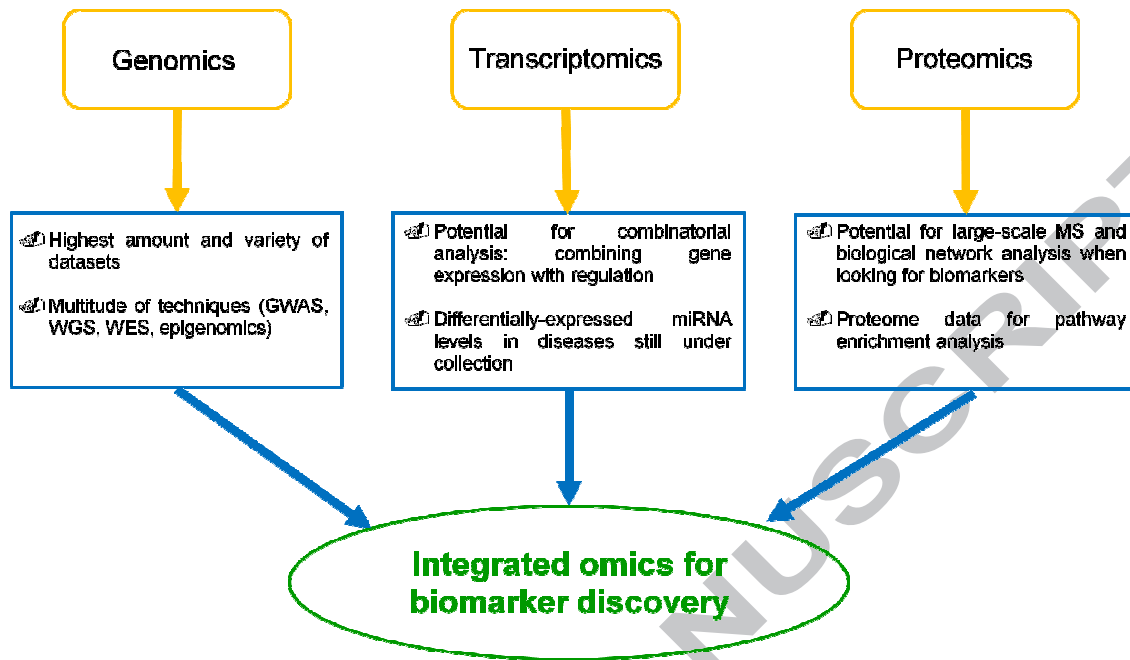
554



555

556

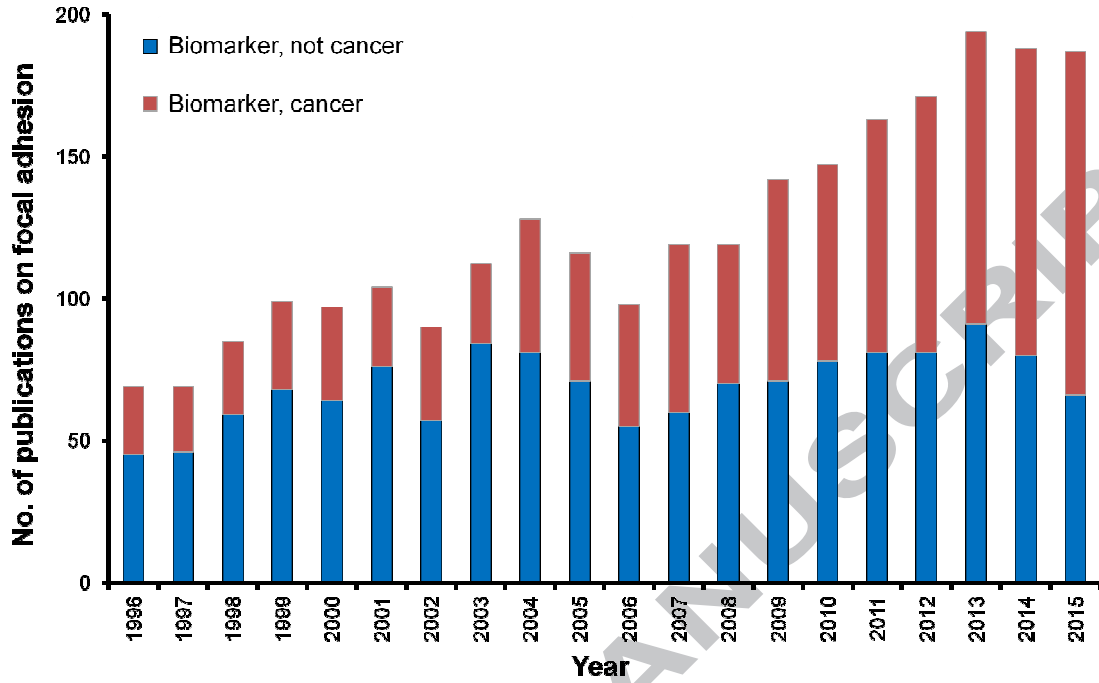
557



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, phosphatidylinositol-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