Charting the molecular links between driver and susceptibility genes in colorectal cancer

Rodrigo Arroyo [a], Miquel Duran-Frigola [a], Clara Berenguer [a], Montserrat Soler-López [a], Patrick Aloy [a,b,*]

[a] Joint IRB-BSC Program in Computational Biology, Institute for Research in Biomedicine (IRB Barcelona), c/Baldiri Reixac 10-12, Barcelona 08028, Spain
[b] Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain

ARTICLE INFO

ABSTRACT

Despite significant advances in the identification of specific genes and pathways important in the onset and progression of colorectal cancer (CRC), mechanistic insight into the relationship between driver and susceptibility genes is needed. In this paper, we systematically explore physical interactions between causative and putative CRC susceptibility genes to reveal the molecular mechanisms involved in tumor biology. In total, we identify 622 high-confidence protein–protein interactions between 42 CRC causative and 65 candidate susceptibility genes. Among the latter, 28 are located in the CRCS9 loci, related to the etiology of CRC, and 17 are co-expressed with well-established CRC drivers, which makes them excellent candidates for further functional studies. Moreover, we find a high degree of functional coherence between connected driver and susceptibility genes, which indicates that our network-based strategy is useful to gain insight into the underlying mechanisms of those proteins with unknown roles in CRC.

1. Introduction

Colorectal cancer (CRC) is a major cause of cancer mortality, with over 600,000 recorded deaths in 2008 [1]. It is the third most common cancer worldwide and the fourth in number of caused deaths, after lung, stomach, and liver cancers [2]. The etiological factors and pathogenetic mechanisms underlying CRC development are complex and heterogeneous, and have been traditionally classified into three types: sporadic (representing 70%), familial (more than 25%) and inherited (less than 10%) [3].

There has been a significant advance in identifying the specific driver genes and pathways important in the initiation and progression of CRC, as well as the constellation of somatic alterations that are present in sporadic CRCs, which include the WNT, RAS-MAPK, PI3K, TGF-β, P53 and DNA mismatch-repair pathways [4]. Furthermore, large-scale sequencing and expression analyses have identified numerous genes recurrently mutated or whose expression is dysregulated in colorectal tumors [5,6]. However, it is not clear what is the relationship between these altered genes and CRC drivers, and further mechanistic insights into these relationships may enable deeper understanding of the pathophysiology of CRC and help identifying novel therapeutic targets.

Recent studies have shown that disease-causing genes for many disease phenotypes often work together within the same biological module [7], be it a protein complex, a pathway or a protein interaction sub-network, highlighting a strong link between protein connectivity and disease [8,9]. Indeed, protein connectivity has been used as a powerful property for the identification of novel relevant genes in complex diseases such as breast cancer [10], Huntington's [11], schizophrenia [12], cerebral ataxias [13] or Alzheimer's [14].

Accordingly, in this paper, we set up an interaction discovery strategy to unveil novel interactions between CRC causative (driver) and putative susceptibility genes, with the aim to provide mechanistic details and understand their relatedness.

2. Materials and methods

2.1. Subcloning of human cDNAs into Y2H plasmids

55 Clones from the human ORFeome v1.1 [15], 27 from Life Technologies Ultimate™ ORF Clones [16] and 5 from the Human ORFeome Collaboration Clone were obtained as a Gateway® cloning adapted plasmids and sequence verified. 3 cDNA clones from human verified full length cDNA repository (ImaGenes) were amplified by PCR with primers containing additional nucleotides (CACC). These cDNAs were cloned into a pENTR™/D-TOPO® vector (pENTR Directional TOPO cloning kit; Life Technologies) and sequenced.

All 90 ORFs were individually transferred into yeast two-hybrid (Y2H) destination vectors by Gateway© recombinational cloning (ProQuest System, Life Technologies). Seed genes were cloned into

pDEST32 to generate bait plasmids and candidate genes were cloned into pDEST22 to obtain prey plasmids.

2.2. Y2H co-transformation screens

Bait and prey plasmids were pair-wise co-transformed into MaV203 yeast strain in 96-well arrays, plated onto selective SD2 (lacking Leu and Trp amino acids) agar media and incubated for 48 h at 30 °C to detect colony growth. Co-transformant arrays were then replica plated onto different selective media for interaction screening. To assay the activation of the HIS3 reporter gene, SD3 (lacking Leu, Trp and His) agar plates were supplemented with 12–100 mM of 3-aminotriazole (3AT, Sigma–Aldrich), being 50 mM 3AT the optimal concentration for positive HIS3 activation colonies. Similarly, we assayed the activation of the URA3 reporter gene by plating onto SD3 (lacking Leu, Trp and Ura). Double reporter HIS3/URA3 activation was evaluated by SD4 (lacking Leu, Trp, His and Ura) agar plates supplemented with 20 mM of 3AT and LacZ reporter gene was tested by the β-galactosidase assay on a nylon membrane placed onto a SD2 agar plate. All configurations were tested in duplicate.

2.3. Y2H library screens

Y2H library screens were performed using an adult human brain cDNA prey library (ProQuest, Life Technologies). The overlapping transcriptome ratio between brain and colon/rectum tissues is 0.93, according to the TissueInfo database [17]. Yeast cells expressing individual baits were transformed with the cDNA library using the LiAc/SS carrier DNA/PEG method [18] and were screened onto selective agar media for reporter gene activation. We checked HIS3 and URA3 reporter gene activation by 7 day incubation of transformed cells at 30 °C in selective agar media (HIS- and/or URA-). We picked up positive growing colonies and cultured them in prey selective liquid medium (lacking Trp). In each screen, approximately 5×10^5 auxotrophic transformants were tested on selective plates, obtaining 20–40 positive colonies on average. The prey plasmid DNA was extracted from the cultures and the bacterial transformation of each plasmid was carried out in order to enable DNA sequencing and subsequent gene identification by BLAST search. The preys identified by the library were tested with their respective baits for activation of reporter gene expression in co-transformation assays, in a similar procedure as explained above.

2.4. Correlation in gene expression profiles

We used microarray data from [19], a compendium of gene expression profiles from 73 normal tissue and cell types. As input, we used this compendium of gene expression profiles, a list of known disease-related genes (OMIM, COSMIC and literature based seed genes) and a list of candidate disease genes. We then applied a mixture model in order to obtain correlation coefficients that were robust to the presence of noise and we fit the model using the Expectation-Maximization (EM) algorithm [20]. The procedure computes gene expression correlation coefficient among disease genes and also between known and candidate disease genes. We defined that known disease gene and a candidate disease gene as co-expressed if their EM correlation coefficient was greater than 2 and the probability of noise less than 2.

2.5. Functional coherence of the HC interactions

Functional similarity between pairs of proteins was computed using best-match average of G-SESAME semantic similarities among annotated GO biological processes [21] [UniProt annotation file (April 30, 2013), including electronically inferred terms; GO ontology file (May 6, 2013), with "cross-products" and "has part" relationships removed]. For each seed and all of its newly identified interactors, we defined functional coherence as the average of pair-wise functional similarities. The significance of this measure was assessed after assigning for 10,000 times an equal number of random proteins from the human binary interactome [22]. Analogously, we calculated a global functional coherence degree by simply averaging the functional similarities of our experimental interactome. To supply compelling evidence that functional coherence around seeds can indeed provide mechanistic insights, we conducted the same procedure with seed direct interactors that are annotated in Reactome pathways (May 6, 2013) [23].

3. Results and discussion

It has been described that, in complex diseases, causative and susceptibility genes tend to be highly interconnected [7]. Based on this observation, we sought to exploit this high interconnectivity to reveal novel direct relationships between CRC causative genes, namely drivers, and a set of putative CRC susceptibility genes. We propose that, by discovering interactions between drivers and susceptibility genes, we can help unraveling the role of the latter in the CRC phenotype and provide mechanistic details for such relationships.

To this end, we compiled a comprehensive list of 45 driver genes. Eleven of these drivers were manually selected for being well-established and characterized genes in CRC development [24] and/or being key genes in the main signaling pathways involved in colorectal cancer syndromes [25]. In addition, we included 24 genes whose defects can strongly contribute to an inherited predisposition to CRC, according to OMIM database [26]. In particular, mutations in any of those 24 genes result in a marked predisposition to colorectal cancer in any of the 2 distinct syndromes: familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC). Finally, by examining COSMIC database [27,28], we selected another 10 genes, whose somatic alterations are known to be crucial in sporadic CRCs. A complete list of the selected driver genes is given in Table S1A.

There is a large number of genes potentially involved in CRC, which have been associated with CRC by several assays such as expression arrays or gene linkage analyses and genome-wide association studies, but whose role in the disease is barely understood. Therefore, in order to reduce the number of putative genes and maximize the biological relevance of the discovered interactions, we selected 45 susceptibility genes after applying the following criteria. First, we prioritized genes with coordinated expression changes with CRC drivers across a compendium of normal tissues and cell types. We estimated co-expression in terms of correlation coefficients computed using an Expectation-Maximization EM algorithm [20]. Besides, in all the cases, we forced a co-expression in the colon and rectum related tissues to obtain the most relevant correlation for CRC (see Section 2). This procedure identified 17 candidate genes that did significantly co-express with the known CRC-causative genes. In addition, we considered genes within chromosomal loci identified by gene linkage analyses and genome-wide association studies as related to CRC. In particular, we included the region 9q22.32-31.1, which contains a susceptibility locus (CRCS9) involved in the development of known hereditary colorectal cancer syndromes, familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC) [29–31]. It is known that these loci are linked to CRC, but the specific genes responsible for this linkage remain to be discovered. Accordingly, we selected 28 additional genes from this locus

potentially implicated in CRC disease mechanisms. The final list of 45 candidate CRC susceptibility genes is given in Table S1B.

We then performed systematic matrix-based yeast two-hybrid (Y2H) screens to identify and characterize novel interactions between driver and susceptibility CRC genes. We converted 45 driver genes into bait and 45 susceptibility genes into prey plasmids, respectively (see Section 2). From those, 3 baits were discarded as they showed self-activation for the four reporter genes in Y2H screens in presence of empty prey clones (BUB1B, CTNNB1 and TP53). Overall, we examined a matrix of 1890 protein pairs by Y2H and identified 1029 non-redundant interactions, which were subsequently scored based on their ability to activate at least two reporter genes or on being observed in the two biological replica screens. We finally selected 595 interactions, which we defined as our Y2H matrix high-confidence (HC) interaction set (Table S1C), and it represents the first molecular links reported between these 45 susceptibility genes and well-understood CRC driver genes.

Although primarily designed to detect a physical association between two known proteins, the Y2H assay can be also used as a system to screen libraries. In order to expand the initial CRC interactome, we carried out an alternative identification of novel CRC-related genes by Y2H library assays, which enable us to identify proteins that bind to a given bait protein in large cDNA libraries. As baits, we selected 6 genes based on gene expression profiles from normal mucosa, colorectal adenoma and adenocarcinoma. In particular, we focused our study on the driver genes AXIN2,

DLC1 and PDGFRL, and the putative susceptibility genes C9orf30, SFRP4 and SFRP2, dysregulated in colorectal tumors compared to adenoma (benign tumor) or normal mucosa [32–34]. Hence, these alterations suggested that these might be specific genes for CRC tumorigenesis. After converting these genes into bait plasmids, we screened them against a human cDNA library through Y2H library assays. We carried out 30 Y2H screens against an adult brain cDNA prey library (5 independent replicates for each of the 6 baits), which yielded 105 interactions between the 6 baits and 74 distinct cDNA clones or preys. DNA sequence verification and a systematic BLAST search showed that 33 of the isolated potential interactors (i.e. preys) contained the downstream gene in frame with GAL4 activation domain, while the remaining clones showed out-of-frame sequences or sequences from non-coding regions, which were discarded. These 33 potential interactors were re-tested by matrix Y2H experiments to validate their interaction specificity. We could validate 27 of them, involving 20 novel proteins, which we then considered as high-confidence interactions and further included in the CRC interactome (Table S1D). Most of the identified preys interacted with a single bait, and only PFKM, RANGAP1 and TMEM129 interacted with two or more baits.

Based on the outcome derived from both Y2H screens, we generated a definitive HC protein–protein interaction network comprising 622 non-redundant interactions, involving 42 of the selected CRC driver genes (Fig. 1). Interestingly, we did not identify common protein–protein interactions in both matrix and library
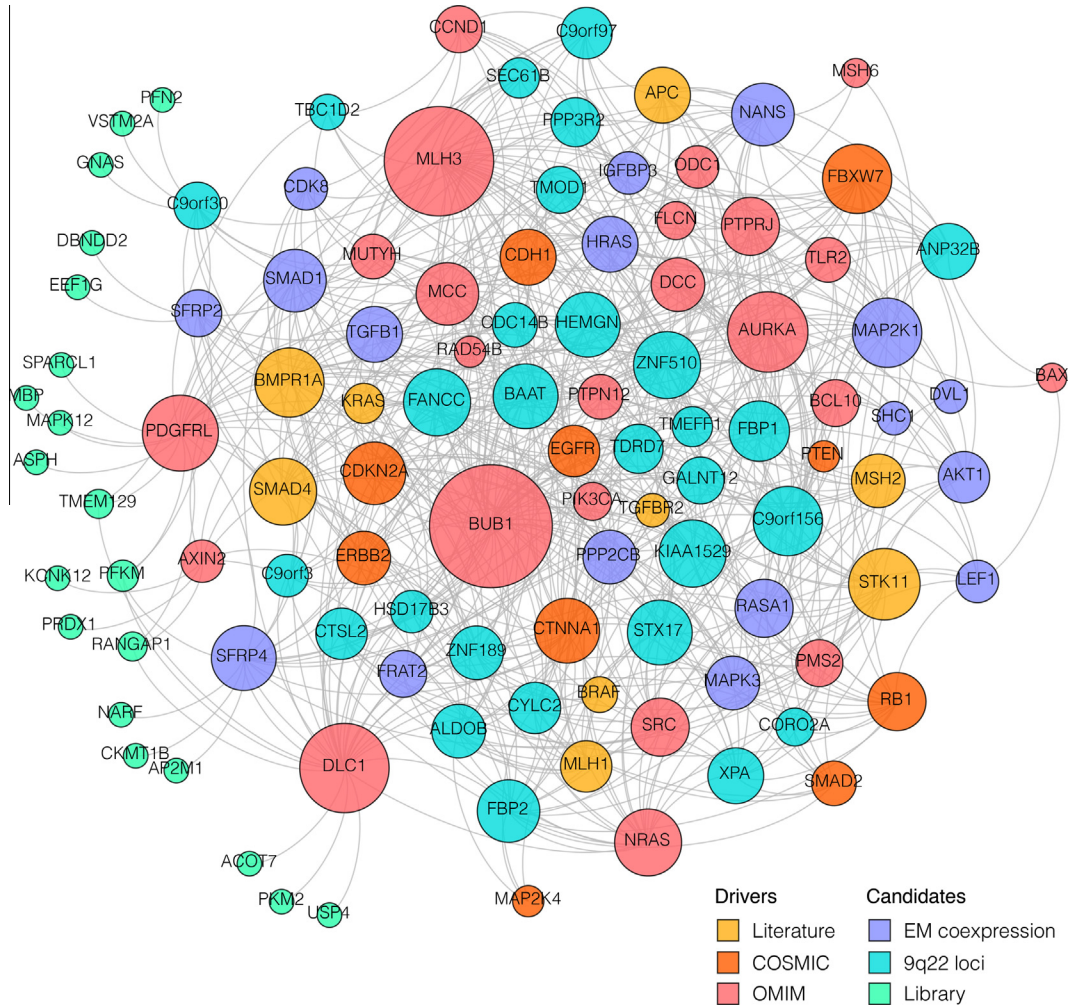

Fig. 1. CRC interaction network. The network comprises 622 interactions involving 42 driver and 65 susceptibility gene products. Drivers are depicted as yellow (literature curated), orange (COSMIC database) and red (OMIM database) nodes, whereas CRC-associated genes are colored in purple (co-expression correlation), blue (CRCS9 loci) and green (library interactors). The size of each node is proportional to its number of interactions.
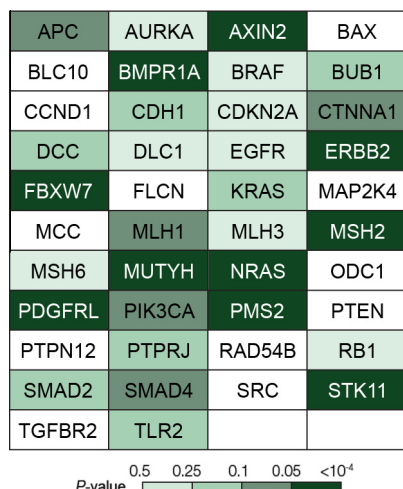
APC	AURKA	AXIN2	BAX
BLC10	BMPR1A	BRAF	BUB1
CCND1	CDH1	CDKN2A	CTNNA1
DCC	DLC1	EGFR	ERBB2
FBXW7	FLCN	KRAS	MAP2K4
MCC	MLH1	MLH3	MSH2
MSH6	MUTYH	NRAS	ODC1
PDGFRL	PIK3CA	PMS2	PTEN
PTPN12	PTPRJ	RAD54B	RB1
SMAD2	SMAD4	SRC	STK11
TGFBR2	TLR2		

P-value 0.5 0.25 0.1 0.05 <10⁻⁴

Fig. 2. Functional coherence of CRC interaction network. The functional coherence of each CRC driver with its interactors is color-graded, from a global tendency to significant coherence.

screens, which reveals the advantage of performing pair-wise and pool screens in parallel. It is worth mentioning that only 6 interactions have been reported since we started the study [22], which indicates that almost all of our HC interactions are novel, as expected. The low overlap between different networks is a well-known effect, and it is mainly attributed to the limited sampling of the interactome space [35,36]. In our case, this is particularly pronounced since we chose some of our candidate genes to maximize the number of novel interactions added to the CRC associated network and of which little interaction information was known (i.e. genes in the susceptibility region CRCS9 with no direct proof of their implication in CRC). The protein interactions reported here have been submitted to the International Molecular Exchange (IMEx) Consortium (http://www.imexconsortium.org) through MINT [37] and assigned the identifier IM-21669.

To evaluate whether the discovered interactions could really provide mechanistic details about the relationships between driver and susceptibility genes, we assessed the functional coherence of our HC interaction network by checking if drivers and their interacting partners are generally involved in similar biological processes, as assigned by Gene Ontology terms. Interestingly, the network shows a high degree of biological process coherence (pairwise semantic similarity mean value = 0.321, P-value < 0.0001) (Fig. 2). The same analysis with driver direct interactors from canonical pathways also shows a high coherence (pairwise semantic similarity mean value = 0.610, P-value < 0.0001), as expected. These findings indicate that our network-based strategy is useful to gain insight into the underlying mechanisms of those proteins with previously unknown roles in CRC, as well as to have a better understanding of the regulation and interrelationship between different proteins of complex biological systems.

4. Concluding remarks

Network and systems biology strategies offer a global perspective to explore the molecular mechanisms that underlie complex diseases beyond individual genes and proteins. In this work, we have shown how an interaction discovery strategy can provide further evidence for potential causative/susceptibility genes related to colorectal cancer. However, to be most valuable, these hints coming from global analyses will need to be individually validated. The finding that causative genes are often highly interconnected in complex heterogeneous disorders, places network biology to

the point where it can be effectively applied to biomedicine in the quest for novel genes associated with human pathologies.

Author contribution

Author contributions: C.B., M.S-L. and P.A. designed research; R.A. and C.B. performed experiments; R.A., M.D-F. and P.A. analyzed data; R.A., M.D-F., M.S-L. and P.A. wrote the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2013.12.012.

References

[1] J. Ferlay, H.R. Shin, F. Bray, D. Forman, C. Mathers, D.M. Parkin, Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008, Int. J. Cancer 127 (2008) 2893–2917.
[2] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA Cancer J. Clin. 61 (2011) 69–90.
[3] J. Souglakos, Genetic alterations in sporadic and hereditary colorectal cancer: implementations for screening and follow-up, Dig. Dis. 25 (2007) 9–19.
[4] E.R. Fearon, Molecular genetics of colorectal cancer, Annu. Rev. Pathol. 6 (2011) 479–507.
[5] T. Sjoblom, S. Jones, L.D. Wood, D.W. Parsons, J. Lin, T.D. Barber, D. Mandelker, R.J. Leary, J. Ptak, N. Silliman, S. Szabo, P. Buckhaults, C. Farrell, P. Meeh, S.D. Markowitz, J. Willis, D. Dawson, J.K. Willson, A.F. Gazdar, J. Hartigan, L. Wu, C. Liu, G. Parmigiani, B.H. Park, K.E. Bachman, N. Papadopoulos, B. Vogelstein, K.W. Kinzler, V.E. Velculescu, The consensus coding sequences of human breast and colorectal cancers, Science 314 (2006) 268–274.
[6] L.D. Wood, D.W. Parsons, S. Jones, J. Lin, T. Sjoblom, R.J. Leary, D. Shen, S.M. Boca, T. Barber, J. Ptak, N. Silliman, S. Szabo, Z. Dezso, V. Ustyanksky, T. Nikolskaya, Y. Nikolsky, R. Karchin, P.A. Wilson, J.S. Kaminker, Z. Zhang, R. Croshaw, J. Willis, D. Dawson, M. Shipitsin, J.K. Willson, S. Sukumar, K. Polyak, B.H. Park, C.L. Pethiyagoda, P.V. Pant, D.G. Ballinger, A.B. Sparks, J. Hartigan, D.R. Smith, E. Suh, N. Papadopoulos, P. Buckhaults, S.D. Markowitz, G. Parmigiani, K.W. Kinzler, V.E. Velculescu, B. Vogelstein, The genomic landscapes of human breast and colorectal cancers, Science 318 (2007) 1108–1113.
[7] M. Oti, H.G. Brunner, The modular nature of genetic diseases, Clin. Genet. 71 (2007) 1–11.
[8] A. Zanzoni, M. Soler-López, P. Aloy, A network medicine approach to human disease, FEBS Lett. 583 (2009) 1759–1765.
[9] A. Pujol, R. Mosca, J. Farres, P. Aloy, Unveiling the role of network and systems biology in drug discovery, Trends Pharmacol. Sci. 31 (2010) 115–123.
[10] M.A. Pujana, J.D. Han, L.M. Starita, K.N. Stevens, M. Tewari, J.S. Ahn, G. Rennert, V. Moreno, T. Kirchhoff, B. Gold, V. Assmann, W.M. Elshamy, J.F. Rual, D. Levine, L.S. Rozek, R.S. Gelman, K.C. Gunsalus, R.A. Greenberg, B. Sobhian, N. Bertin, K. Venkatesan, N. Ayivi-Guedehoussou, X. Sole, P. Hernandez, C. Lazaro, K.L. Nathanson, B.L. Weber, M.E. Cusick, D.E. Hill, K. Offit, D.M. Livingston, S.B. Gruber, J.D. Parvin, M. Vidal, Network modeling links breast cancer susceptibility and centrosome dysfunction, Nat. Genet. 39 (2007) 1338–1349.
[11] H. Goehler, M. Lalowski, U. Stelzl, S. Waelter, M. Stroedicke, U. Worm, A. Droege, K.S. Lindenberg, M. Knoblich, C. Haenig, M. Herbst, J. Suopanki, E. Scherzinger, C. Abraham, B. Bauer, R. Hasenbank, A. Fritzsche, A.H. Ludewig, K. Bussow, S.H. Coleman, C.A. Gutekunst, B.G. Landwehrmeyer, H. Lehrach, E.E. Wanker, A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease, Mol. Cell 15 (2004) 853–865.
[12] L.M. Camargo, V. Collura, J.C. Rain, K. Mizuguchi, H. Hermjakob, S. Kerrien, T.P. Bonnert, P.J. Whiting, N.J. Brandon, Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia, Mol. Psychiatry 12 (2007) 74–86.
[13] J. Lim, T. Hao, C. Shaw, A.J. Patel, G. Szabo, J.F. Rual, C.J. Fisk, N. Li, A. Smolyar, D.E. Hill, A.L. Barabasi, M. Vidal, H.Y. Zoghbi, A protein–protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration, Cell 125 (2006) 801–814.

[14] M. Soler-Lopez, A. Zanzoni, R. Lluis, U. Stelzl, P. Aloy, Interactome mapping suggests new mechanistic details underlying Alzheimer's disease, Genome Res. 21 (2011) 364–376.

[15] P. Lamesch, N. Li, S. Milstein, C. Fan, T. Hao, G. Szabo, Z. Hu, K. Venkatesan, G. Bethel, P. Martin, J. Rogers, S. Lawlor, S. McLaren, A. Dricot, H. Borick, M.E. Cusick, J. Vandenhaute, I. Dunham, D.E. Hill, M. Vidal, hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes, Genomics 89 (2007) 307–315.

[16] F. Liang, U. Matrubutham, B. Parvizi, J. Yen, D. Duan, J. Mirchandani, S. Hashima, U. Nguyen, E. Ubil, J. Loewenheim, X. Yu, S. Sipes, W. Williams, L. Wang, R. Bennett, J. Carrino, ORFDB: an information resource linking scientific content to a high-quality Open Reading Frame (ORF) collection, Nucleic Acids Res. 32 (2004) D595–D599.

[17] L. Skrabanek, F. Campagne, TissueInfo: high-throughput identification of tissue expression profiles and specificity, Nucleic Acids Res. 29 (2001) E102.

[18] R.D. Gietz, R.H. Schiestl, Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method, Nat. Protoc. 2 (2007) 38–41.

[19] A.I. Su, T. Wiltshire, S. Batalov, H. Lapp, K.A. Ching, D. Block, J. Zhang, R. Soden, M. Hayakawa, G. Kreiman, M.P. Cooke, J.R. Walker, J.B. Hogenesch, A gene atlas of the mouse and human protein-encoding transcriptomes, Proc. Natl. Acad. Sci. U S A 101 (2004) 6062–6067.

[20] A. Dempster, N. Laird, D. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, J. R. Statist. Soc. Ser. B Methodol. 39 (1977) 1–38.

[21] Z. Du, L. Li, C.F. Chen, P.S. Yu, J.Z. Wang, G-SESAME: web tools for GO-term-based gene similarity analysis and knowledge discovery, Nucleic Acids Res. 37 (2009) W345–W349.

[22] R. Mosca, A. Céol, P. Aloy, Interactome3D: adding structural details to protein networks, Nat. Methods 10 (2013) 47–53.

[23] L. Matthews, G. Gopinath, M. Gillespie, M. Caudy, D. Croft, B. de Bono, P. Garapati, J. Hemish, H. Hermjakob, B. Jassal, A. Kanapin, S. Lewis, S. Mahajan, B. May, E. Schmidt, I. Vastrik, G. Wu, E. Birney, L. Stein, P. D'Eustachio, Reactome knowledgebase of human biological pathways and processes, Nucleic Acids Res. 37 (2009) D619–D622.

[24] S.D. Markowitz, M.M. Bertagnolli, Molecular origins of cancer: molecular basis of colorectal cancer, N. Engl. J. Med. 361 (2009) 2449–2460.

[25] E. Sancho, E. Batlle, H. Clevers, Signaling pathways in intestinal development and cancer, Annu. Rev. Cell Dev. Biol. 20 (2004) 695–723.

[26] V.A. McKusick, Mendelian Inheritance in Man and its online version, OMIM, Am. J. Hum. Genet. 80 (2007) 588–604.

[27] S. Bamford, E. Dawson, S. Forbes, J. Clements, R. Pettett, A. Dogan, A. Flanagan, J. Teague, P.A. Futreal, M.R. Stratton, R. Wooster, The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website, Br. J. Cancer 91 (2004) 355–358.

[28] S.A. Forbes, G. Tang, N. Bindal, S. Bamford, E. Dawson, C. Cole, C.Y. Kok, M. Jia, R. Ewing, A. Menzies, J.W. Teague, M.R. Stratton, P.A. Futreal, COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer, Nucleic Acids Res. 38 (2010) D652–D657.

[29] G.L. Wiesner, D. Daley, S. Lewis, C. Ticknor, P. Platzer, J. Lutterbaugh, M. MacMillen, B. Baliner, J. Willis, R.C. Elston, S.D. Markowitz, A subset of familial colorectal neoplasia kindreds linked to chromosome 9q22.2-31.2, Proc. Natl. Acad. Sci. U S A 100 (2003) 12961–12965.

[30] Z.E. Kemp, L.G. Carvajal-Carmona, E. Barclay, M. Gorman, L. Martin, W. Wood, A. Rowan, C. Donohue, S. Spain, E. Jaeger, D.G. Evans, E.R. Maher, T. Bishop, H. Thomas, R. Houlston, I. Tomlinson, Evidence of linkage to chromosome 9q22.33 in colorectal cancer kindreds from the United Kingdom, Cancer Res. 66 (2006) 5003–5006.

[31] J. Skoglund, T. Djureinovic, X.L. Zhou, J. Vandrovcova, E. Renkonen, L. Iselius, M.L. Bisgaard, P. Peltomaki, A. Lindblom, Linkage analysis in a large Swedish family supports the presence of a susceptibility locus for adenoma and colorectal cancer on chromosome 9q22.32-31.1, J. Med. Genet. 43 (2006) e7.

[32] J. Sabates-Bellver, L.G. Van der Flier, M. de Palo, E. Cattaneo, C. Maake, H. Rehrauer, E. Laczko, M.A. Kurowski, J.M. Bujnicki, M. Menigatti, J. Luz, T.V. Ranalli, V. Gomes, A. Pastorelli, R. Faggiani, M. Anti, J. Jiricny, H. Clevers, G. Marra, Transcriptome profile of human colorectal adenomas, Mol. Cancer Res. 5 (2007) 1263–1275.

[33] L.G. Van der Flier, J. Sabates-Bellver, I. Oving, A. Haegebarth, M. De Palo, M. Anti, M.E. Van Gijn, S. Suijkerbuijk, M. Van de Wetering, G. Marra, H. Clevers, The intestinal Wnt/TCF signature, Gastroenterology 132 (2007) 628–632.

[34] C.B. Diep, M.R. Teixeira, L. Thorstensen, J.N. Wiig, M. Eknaes, J.M. Nesland, K.E. Giercksky, B. Johansson, R.A. Lothe, Genome characteristics of primary carcinomas, local recurrences, carcinomatoses, and liver metastases from colorectal cancer patients, Mol. Cancer 3 (2004) 6.

[35] K. Venkatesan, J.F. Rual, A. Vazquez, U. Stelzl, I. Lemmens, T. Hirozane-Kishikawa, T. Hao, M. Zenkner, X. Xin, K.I. Goh, M.A. Yildirim, N. Simonis, K. Heinzmann, F. Gebreab, J.M. Sahalie, S. Cevik, C. Simon, A.S. de Smet, E. Dann, A. Smolyar, A. Vinayagam, H. Yu, D. Szeto, H. Borick, A. Dricot, N. Klitgord, R.R. Murray, C. Lin, M. Lalowski, J. Timm, K. Rau, C. Boone, P. Braun, M.E. Cusick, F.P. Roth, D.E. Hill, J. Tavernier, E.E. Wanker, A.L. Barabasi, M. Vidal, An empirical framework for binary interactome mapping, Nat. Methods 6 (2009) 83–90.

[36] R.B. Russell, P. Aloy, Targeting and tinkering with interaction networks, Nat. Chem. Biol. 4 (2008) 666–673.

[37] A. Chatr-aryamontri, A. Ceol, L.M. Palazzi, G. Nardelli, M.V. Schneider, L. Castagnoli, G. Cesareni, MINT: the Molecular INTeraction database, Nucleic Acids Res. 35 (2007) D572–D574.