Clin Transl Oncol (2013) 15:3–8
DOI 10.1007/s12094-012-0922-z

Next-generation sequencing reveals the secrets of the chronic lymphocytic leukemia genome

Andrew J. Ramsay · Alejandra Martínez-Trillos · Pedro Jares · David Rodríguez · Agnieszka Kwarciak · Víctor Quesada

Received: 29 June 2012 / Accepted: 23 July 2012 / Published online: 22 August 2012

Abstract The study of the detailed molecular history of cancer development is one of the most promising techniques to understand and fight this diverse and prevalent disease. Unfortunately, this history is as diverse as cancer itself. Therefore, even with next-generation sequencing techniques, it is not easy to distinguish significant (driver) from random (passenger) events. The International Cancer Genome Consortium (ICGC) was formed to solve this fundamental issue by coordinating the sequencing of samples from 50 different cancer types and/or sub-types that are of clinical and societal importance. The contribution of Spain in this consortium has been focused on chronic lymphocytic leukemia (CLL). This approach has unveiled new and unexpected events in the development of CLL. In this review, we introduce the approaches utilized by the consortium for the study of the CLL genome and discuss the recent results and future perspectives of this work.

Keywords Chronic lymphocytic leukemia · Cancer genome · Next-generation sequencing

A. J. Ramsay · D. Rodríguez · A. Kwarciak · V. Quesada (✉)
Departamento de Bioquímica y Biología Molecular, Instituto Universitario de Oncología, IUOPA, Universidad de Oviedo, 33006 Oviedo, Spain
e-mail: quesadavictor@uniovi.es

A. Martínez-Trillos · P. Jares
Unidad de Hematopatología, Servicio de Anatomía Patológica, Hospital Clínic, Universitat de Barcelona, IDIBAPS, Barcelona, Spain

The genomes of cancer

Cancer is a genetic disease arising from unregulated clonal expansion of distinct cell types [1]. The cancer genome is the unique genetic profile of cancer cells acquired through accumulative genetic alterations [2]. The characteristic genetic alterations—including chromosomal rearrangements, copy number changes and point mutations—and epigenetic changes—e.g. DNA methylation and histone modification—manifest somatically in the cancer cells, and are not present in the germline. An individual's repertoire of germline variants (present in both normal and cancerous cells) can influence cancer susceptibility and progression by altering important cellular features in cancerous clones, such as growth and metabolism. Somatic mutations occur in all cells during their lifespan and manifest randomly throughout the genome. The acquisition rates of these mutations can be enhanced through DNA damage or the failure of DNA repair machinery [1]. Individual tumors host a spectrum of somatic genomic and epigenetic abnormalities that can vary significantly between clinically similar cancer types. A subset of mutations that provide the cancerous cells with advantageous survival phenotypes—termed driver mutations—occur frequently in multiple cancers and may define cancer subtypes [3]. Mutations that do not constitute the compendium of driver mutations within cancerous cells are termed passenger mutations, and can be acquired coincidentally or before driving mutations. As an illustrative point, it is known that solid tumors typically contain between 40 and 100 coding gene alterations, of which just 5–15 are classified as driver mutations [4]. Identifying and cataloging such driver mutations from the genomes of diverse cancers has long been a focal point for researchers and has been instrumental for our current understanding of the fundamental molecular aspects of

cancer. The International Cancer Genome Consortium (ICGC) was formed to coordinate large-scale cancer genome studies utilizing next-generation sequencing (NGS) in tumors from 50 different cancer types and/or subtypes that are of clinical and societal importance across the globe [5]. Spain's own contribution to the ICGC has been the study of the chronic lymphocytic leukemia (CLL) genome. This landmark project has so far had significant success in identifying the repertoire of oncogenic mutations that drive CLL and will potentially enable the development of new cancer therapies. In this review, we introduce next-generation sequencing and overview the approaches utilized by the consortium for the study of the CLL genome and highlight the recent successes and future perspectives of this work.

The chronic lymphocytic leukemia genome

The draft of the human genome in 2,000 culminated the achievements of first generation sequencing (Sanger's sequencing) and paved the way for the discovery of many important driver mutations in the cancer genome. However, this approach is limited both in resolution and throughput, hampering the examination of non-coding regions and the analysis of large patient sample cohorts. The emergence of next-generation sequencing (NGS) offers higher throughput and greatly increased sensitivity, enabling the sequencing of complete genomes, exomes (all annotated exons) or transcriptomes (all RNA transcripts) of individual cancers. Whole cancer genome sequencing is the complete sequencing of the genome of cancer tissue using the germline genome as a control [6, 7]. The most defining characteristic of whole-genome sequencing, and indeed next-generation sequencing as a whole, is massive parallelization, i.e. the ability to obtain millions of sequences in a single experiment. This allows the collection of sequence information at any nucleotide position in the genome with high redundancy, which in turn, permits the identification of frequent and infrequent genome changes in heterogeneous cancer samples and can discriminate the entire range of genomic alterations in a single experiment. This then makes whole-genome sequencing of cancer genomes the most comprehensive of the NGS approaches. The first whole cancer genome sequence was reported in 2008 from an acute myeloid leukemia (AML), and demonstrated the power of the approach for discovering novel somatic mutations [8]. A rapidly growing number of cancer genomes have been characterized using whole-genome sequencing or combinatorial application of exome sequencing, transcriptome sequencing and structural variation analysis. At present, these cancer genomes include, AML [8], lung [9], multiple myeloma [10], breast [11], melanoma [12], pancreatic [13], liver [14], medulloblastomas [15], prostate [16], acute promyleolocytic leukemia [17], pediatric diffuse intrinsic pontine gliomas [18] ovarian [19] and chronic lymphocytic leukemia [20–23].

Chronic lymphocytic leukemia (CLL) is a common neoplasia of B lymphocytes that is characterized by the accumulation of CD5-positive monoclonal B cells in blood, bone marrow and peripheral lymphoid organs [24, 25]. CLL exhibits a heterogeneous clinical course ranging from an indolent disorder with a normal lifespan, to an aggressive disease and short survival. These clinically distinct CLL subtypes are characterized by high and low number of somatic hypermutations (SHM) in the variable region of the immunoglobulin genes [26]. A long-standing barrier in understanding leukemogenesis has been the accurate identification of the genetic alterations that underlie this disease. The 2011 report of the CLL genome by the Spanish CLL genome consortium was the first in this frequent leukemia [20]. This breakthrough study by Puente et al. reported the whole genome sequencing of two cases of CLL without mutations in immunoglobulin genes (*IGHV*-unmutated) and two with mutations in immunoglobulin genes (*IGHV*-mutated). CLL tumor samples were obtained from blood-derived mononuclear populations that had been depleted of contaminating T cells, natural killer cells, monocytes and granulocytes. Leukocytes derived from whole blood extracts obtained after treatment and complete remission served as matched normal tissue controls. Whole-genome sequencing of the tumors using germline DNA as a baseline identified approximately 1,000 somatic mutations in each of the tumors, consistent with prior estimates of one mutation per Mb for leukemias and signifies that CLL [8, 10], in addition to medulloblastomas [15] carries comparatively less mutations than lung or melanoma cancers [12, 27]. As expected, Puente et al. detected more somatic mutations in the immunoglobulin genes of the two *IGHV*-mutated patients when compared with the *IGHV*-unmutated patients (16 and 6 vs. 4 and 0). From all of the somatic mutations identified in the four patients, 46 mutations were predicted to alter the protein-coding sequence of 45 genes, of which 26 were found to be expressed at the RNA level and of potential biological relevance. Sanger sequencing in a validation cohort of 169 additional CLL patients showed that four of these genes (NOTCH1, MYD88, XPO1 and KLHL6) were recurrently mutated.

NOTCH1 emerged as the most recurrent target of genetic lesions. *NOTCH1* encodes a class I transmembrane protein that serves as a ligand-activated transcription factor implicated in cell differentiation, proliferation, and apoptosis [28]. In response to ligand binding, NOTCH1 is proteolytically cleaved to free the intracellular portion of the protein for translocation to the nucleus, initiating the

transcriptional activation of multiple target genes. *NOTCH1* was mutated in 12 % of the CLL patients examined (31/255). These mutations are all predicted to generate truncated proteins that lack the C-terminal PEST domains and lead to a more stable and active isoform. Gene expression analysis of cells from 10 of the CLL patients with *NOTCH1* mutations and patients without mutations in this gene found a significant upregulation of genes involved in the NOTCH signaling pathway. Of particularly relevance, in T cell acute lymphoblastic leukemia (T-ALL), activating mutations of *NOTCH1* are the predominant genetic alteration, accounting for up to 60 % of the cases [29] and are responsible for the activation of multiple biosynthetic routes [30], including oxidative phosphorylation and glycolysis/gluconeogenesis. These two metabolic pathways were similarly found to be transcriptionally upregulated in *NOTCH1*-mutated CLL tumors. A fraction of CLL is prone to transformation to diffuse large B cell lymphoma (DLBCL), a condition known as Richter syndrome (RS) that is characterized by an adverse clinical course [31]. Interestingly, *NOTCH1*-mutated CLL cases were found to undergo transformation into DLBCL more frequently and were associated with advanced disease at diagnosis and more severe biological features than cases with wild-type *NOTCH1*. In confirmation of these findings, an independent study confirmed mutational activation of *NOTCH1* in 8.3 % of CLL at diagnosis, and was detected at significantly higher frequency during disease progression toward Richter transformation (31.0 %), as well as in chemorefractory CLL (20.8 %) [23]. Intriguingly, Puente et al. found that two of the four mutations of *exportin 1* (*XPO1*) occur in patients with mutations of *NOTCH1*, suggesting that both mutations may have a synergistic effect in CLL development. XPO1 is a nuclear export receptor involved in the active transport of tumor suppressors, such as p53 and nucleophosmin [32, 33]. Excitingly, XPO1 was recently targeted by a novel therapeutic strategy to restore tumor suppressor function in a murine AML model [34].

In addition among the genes found frequently mutated in the CLL genome by Puente et al. activating mutations of *MYD88* were identified in 9 of 310 patients. This protein is a key component of the Toll-like receptor (TLR) signaling pathway [35]. Consistent with this functional role, CLL tumors containing *MYD88* mutations displayed elevated activation of the downstream effectors STAT3 and NFkB p65 subunit, and increased sensitivity towards agonist stimulation of ILR1 or the TLRs. Identical mutations of this gene are present in various lymphomas, which suggests that *MYD88* may be a proto-oncogene important in the pathogenesis of lymphoid neoplasms [36]. Similar to *NOTCH1* mutations, patients with *MYD88* mutations are associated with advanced disease at diagnosis. Finally, Kelch-like protein 6 (KLHL6) has been implicated in the formation of the germinal centre during B-cell maturation and B cell antigen receptor (BCR) signal transduction [37]. In addition to somatic mutation of *NOTCH1*, *MYD88*, *XPO1* and *KLHL6* in the genome of CLL tumors, deletion of 13q14 was detected in three of the four CLL patients. Deletion of 13q14 affecting the DLEU2/MIR15-16 cluster has been suggested to be an initiating event in nearly 50 % of CLL cases [38–40]. In addition, a 40 Mb deletion in chromosome 6q14–q22 was also detected in one patient, and has been reported in 10 % of CLL patients [38].

The initial draft of the CLL genome was rapidly succeeded by a report by the Spanish consortium of the whole-exome sequencing of 105 CLL patients [21]. Exome sequencing—i.e. the targeted sequencing of all coding regions of the genome—facilitated this comprehensive study to expeditiously evolve due to the greater throughput of this technique. In agreement with whole-genome data, approximately 0.9 mutations were identified per Mb in the 105 CLL patients, comprising 60 *IGHV*- and 45 *IGHV*-unmutated cases. A total of 1,246 mutations were located in coding regions of the 105 patients, excluding those encoding *IGHV*, of which a higher mutational content was found in the *IGHV*-mutated CLL than in *IGHV*-unmutated CLL. The somatic mutations were distributed amongst 1,100 genes, with 62 genes classified as frequently mutated when the number of mutations, size and codon composition of the genes were considered. Functional clustering of the mutated genes showed a significant enrichment in genes of pathways involved in inflammatory response, pattern recognition and DNA damage, amongst others. Exomic sequencing of the large CLL patient cohort was successful in identifying genes recurrently mutated and distinct from those previously identified in whole-genome sequencing-based study of CLL patients [20]. These genes included those encoding SF3B1, one of the subunits of the spliceosomal U2 snRNP [41]; POT1, a nuclear protein involved in telomere maintenance [42, 43]; CHD2, which regulates genome stability maintenance [44]; and LRP1B, which has been recently defined as a tumor suppressor in different malignancies [45]. Interestingly, *POT1* somatic mutations were only in *IGHV*-unmutated tumors, whereas *CHD2* somatic mutations exclusively appeared in *IGHV*-mutated tumors, consistent with different mechanisms for the development of indolent and aggressive CLL subtypes. Only one of the CLL cases carried a *TP53* mutation, which is probably due to the low-risk CLL characteristics of a non-selected population-based series rather than patients in referral or clinical trial studies.

Interestingly, *SF3B1* displayed the somatic point mutations p.K700E in four patients and p.N626Y in two patients. In a validation series of 279 paired tumor and normal samples from CLL patients, 27 *SF3B1* somatic

mutations were detected (9.7 %), with three additional recurrently mutated residues: p.T663I, p.K666E and p.G742D. These findings make *SF3B1* one of the most frequently mutated genes reported to date in CLL, along with *NOTCH1*. Clinical analysis showed that patients with *SF3B1* somatic mutations present advanced disease at diagnosis and have significantly shorter time to disease progression and lower 10-year overall survival rates [21, 46]. Cox analyses suggested that the mutational status of *SF3B1* has a prognostic value independent from clinical stage, ZAP-70 or CD38 expression. In subsequent independent CLL studies, *SF3B1* mutations were associated with deletions of chromosomal region 11q22 and ATM mutations (ATM is located at chromosome 11q22), as well as poor prognosis and resistance to fludarabine therapy [22, 47]. Interestingly, somatic mutations were also identified by Quesada et al. in other genes involved in the splicing machinery such as *SFRS1* (p.Y82*, p.G4Gfs*2), *SFRS7* (p.L18Q) and *U2AF2* (p.Q143L, p.Q190L), indicating that alterations in this post-transcriptional pathway may be particularly relevant in CLL pathogenesis. Similarly, eight genes encoding proteins involved in RNA splicing are mutated with a variable frequency in myelodysplastic syndromes (MDS) [48–50]. In addition to *SF3B1*, *U2AF35* and *SRSF2* are among the most frequently mutated in MDS and may be associated with prognosis and phenotype. Splicing is a pleiotropic mechanism necessary for cell functioning, although specific alterations in the splicing of oncogenes and tumor suppressors have been related to cancer development [51]. *SF3B1* encodes a protein involved in the binding of the spliceosomal U2 snRNP to the branch point close to 3′ splicing sites of mRNA [52, 53]. This protein ensures the fidelity of the 3′ branching site, therefore activation of cryptic 3′ splice sites is the expected effect of altering SF3B1 function. Using comparative analysis of exon arrays, Quesada et al. uncovered a set of 184 genes with exons showing differential inclusion levels in *SF3B1* cells. Further, NGS of total RNA (RNA-seq) demonstrated that transcripts with splicing junctions that were differentially expressed between *SF3B1*-mutated and unmutated tumors contained abnormal 3′ acceptor sites. These novel isoforms included truncated versions of

Table 1 Comparison of recurrent mutations identified in independent studies of CLL patients using whole-genome or whole-exome sequencing

Gene	Puente et al./Quesada et al.		Wang et al.		Fabbri et al.		Cosmic*									
	Discovery	Validation	Discovery	Validation	Discovery	Validation	B	E	O	P	GI	M	L	LY	CNS	Other
NOTCH1	**1/4 WG**	**29/255**	**4/91 Ex**	**8/101**[a]	**2/5 WG**	**8/53**	***15***		***3***		***35***	***16***	***16***	***663***	***12***	
MYD88	**1/4 WG; 2/105 Ex**	**9/310**	**9/91 Ex**	**5/101**[b]	**1/5 WG**	–	***1***				***3***			***109***		
SF3B1	**14/105 Ex**	**27/279**	**14/91 Ex**	**3/101**[c]	–	–	***3***	***1***	***1***	***3***	***4***	***2***	***1***	***334***		***4***
XPO1	**1/4 WG; 2/105 Ex**	**4/279**	**1/91 Ex**	–	–	–	***1***		***1***		***6***	***2***		***5***	***1***	
CHD2	**5/105 Ex**	–	**1/91 Ex**	–	–	–	***1***		***1***		***8***	***1***		***5***		
POT1	**5/105 Ex**	–	**1/91 Ex**	–	–	–	***2***		***1***		***3***	***2***		***3***	***1***	***1***
ZMYM3	**2/105 Ex**	–	**4/91 Ex**	–	–	–			***1***		***8***			***2***		
DDX3X	**2/105 Ex**	–	**3/91 Ex**	–	–	–	***3***		***1***	***1***	***4***	***2***		***2***	***10***	
TGM7	–	–	**1/91 Ex**	–	**1/5 WG**	**2/530**			***2***		***4***					
LRP1B	**5/105 Ex**	–	–	–	–	–	***2***		***13***	***6***	***43***	***8***	***16***	***7***	***2***	***3***
KLHL6	**1/4 WG**	**3/279**	–	–	–	–			***2***		***5***			***1***		
MAPK1	–	–	**3/91 Ex**	–	–	–	***1***		***1***		***1***	***2***		***1***	***1***	
FBXW7	–	–	**4/91 Ex**	–	–	–	***5***	***13***	***9***	***3***	***146***		***14***	***124***	***3***	***15***
BIRC3	–	–	–	–	**1/5 WG**	**2/530**			***3***		***2***	***1***				
PLEKHG5	–	–	–	–	**1/5 WG**	**2/530**										***1***

WG whole genome, *Ex* whole exome, *B* breast, *E* endometrial, *O* ovary, *P* pancreas, *GI* gastrointestinal tract, *M* melanoma, *L* lung, *LY* lymphoid neoplasms, *CNS* central nervous system

* Summary of cancer mutations previously identified in cosmic v59 (http://www.sanger.ac.uk/genetics/CGP/cosmic/). Bold mark positive identification in the respective cancer types. Bold italic represent the amount of unique samples with mutations

[a] *NOTCH1* p.P2514 fs mutation only

[b] *MYD88* p.P258L & p.L265P mutations only

[c] *SF3B1* p.K700E mutation only

SLC23A2, encoding a vitamin C transporter [54], *TCIRG1*, one of whose products, TIRC7, is a T cell immune regulator [55] and *FOXP1*, encoding a member of the forkhead transcription factor group, whose altered expression has been linked to diffuse large B cell lymphoma [56, 57].

An independent CLL genome study by researchers at the Broad Institute (MIT, Boston, USA) emerged simultaneous to the whole-exome report by the Spanish CLL consortium. Wang and colleges sequenced DNA samples of leukemia cells from 91 patients with chronic lymphocytic leukemia (88 exomes and 3 genomes), representing the broad clinical spectrum of the disease [22]. The researchers identified nine genes with significant mutational rates: *TP53*, *SF3B1*, *MYD88*, *ATM*, *FBXW7*, *NOTCH1*, *ZMYM3*, *DDX3X*, and *MAPK1*, with validation of the three most frequently recurring mutations—*SF3B1*–K700E, *MYD88*–L265P, and *NOTCH1*–P2514 fs—in 101 independent paired tumor and germline DNA samples. Comparative analysis of the most recurrently mutated genes identified by the whole-genome and exome sequencing studies of Puente et al. Quesada et al., Wang et al. and Fabbri et al. demonstrates, as expected, a significant overlap in *NOTCH1*, *SF3B1* and *MYD88* (Table 1). Further, overlaps in the less frequently mutated genes *XPO1*, *CHD2*, *POT1*, *ZMYM3, DDX3X* and *TGM7* can be observed (Table 1). Several of the overlapping genes do not have prior associations with CLL and warrant future investigation. As such, *DDX3X* encodes an RNA helicase that functions at multiple levels of RNA processing, including RNA splicing and transport, and has been implicated in tumor proliferation and viral infections [57, 58]. Intriguingly, DDX3X interacts directly with XPO1, another frequently mutated gene in CLL, and co-opts nuclear export [59]. A translocation involving ZMYM3 has been previously correlated to X-linked mental retardation [60], and, relevantly, chromosomal translocation between the homologous family member ZMYM2 and fibroblast growth factor receptor 1 (FGFR1) are linked to lymphoblastic lymphoma and a myeloproliferative disorder [61]. Finally, TGM7 is a transglutaminase of undetermined biological function [62].

Future perspectives

The global initiative that is the ICGC now includes large-scale sequencing projects in Australia, Canada, China, France, Germany, India, Italy, Japan, Mexico, Spain, The United Kingdom and The United States. Extensive sequencing of all of the clinically significant cancers will, according to some estimates, total some tens of thousands of sequenced cancer genomes within the next 7 years [63]. The comprehensive large scale description of all cancer genomes is envisaged to provide adequate resolution within each of the respective cancer subtypes, so that even cancer genes that occur at low frequency will be cataloged. One area which will undoubtedly benefit greatly from knowledge of the cancer genome will be cancer diagnostics. The shrinking costs of NGS has led to speculation that the $1000 genome is almost upon us and could enable individual genome sequencing as a routine clinical test for precise diagnosis, prognosis and targeted therapeutic approaches [64]. Indeed, commercial avenues already exist within Spain for clinicians wishing to employ next-generation approaches (http://www.dreamgenics.com and http://www.gatc-biotech.com). It is foreseeable that continual improvement of NGS sensitivity may allow accurate diagnosis from even smaller tumoral sources, such as circulating tumor cells [65], or mutated cancer DNA that has leaked into the blood or body fluids [66]. In conclusion, NGS and the study of cancer genomes have in many aspects revolutionalized our current understanding of cancer biology. The future prospect of 'genomics' transitioning from the bench to the bedside will inevitably prove to be an exciting and beneficial era for modern medicine.

Acknowledgments This work was funded by the Spanish Ministry of Economy and Competitiveness through the Instituto de Salud Carlos III (ISCIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII. We are very grateful to all members of the Spanish CLL Genome Consortium for their continuous support and collaboration. We are also very grateful to all patients with CLL who have participated in this study.

References

1. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 5:646–674
2. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. Nature 7239:719–724
3. Greenman C, Stephens P, Smith R et al (2007) Patterns of somatic mutation in human cancer genomes. Nature 7132:153–158
4. Bozic I, Antal T, Ohtsuki H et al (2010) Accumulation of driver and passenger mutations during tumor progression. Proc Natl Acad Sci USA 43:18545–18550
5. Hudson TJ, Anderson W, Artez A et al (2010) International network of cancer genome projects. Nature 7291:993–998
6. Zhao J, Grant SF (2011) Advances in whole genome sequencing technology. Curr Pharm Biotechnol 2:293–305
7. Meyerson M, Gabriel S, Getz G (2010) Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet 10:685–696
8. Ley TJ, Mardis ER, Ding L et al (2008) DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 7218:66–72
9. Campbell PJ, Stephens PJ, Pleasance ED et al (2008) Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. Nat Genet 6:722–729
10. Chapman MA, Lawrence MS, Keats JJ et al (2011) Initial genome sequencing and analysis of multiple myeloma. Nature 7339:467–472
11. Stephens PJ, McBride DJ, Lin ML et al (2009) Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature 7276:1005–1010
12. Pleasance ED, Cheetham RK, Stephens PJ et al (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 7278:191–196
13. Campbell PJ, Yachida S, Mudie LJ et al (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 7319:1109–1113
14. Totoki Y, Tatsuno K, Yamamoto S et al (2011) High-resolution characterization of a hepatocellular carcinoma genome. Nat Genet 5:464–469
15. Parsons DW, Li M, Zhang X et al (2011) The genetic landscape of the childhood cancer medulloblastoma. Science 6016:435–439

16. Berger MF, Lawrence MS, Demichelis F et al (2011) The genomic complexity of primary human prostate cancer. Nature 7333:214–220
17. Welch JS, Westervelt P, Ding L et al (2011) Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. JAMA 15:1577–1584
18. Wu G, Broniscer A, McEachron TA et al (2012) Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet 3:251–253
19. Network CGAR (2011) Integrated genomic analyses of ovarian carcinoma. Nature 7353:609–615
20. Puente XS, Pinyol M, Quesada V et al (2011) Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 7354:101–105
21. Quesada V, Conde L, Villamor N et al (2012) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet 1:47–52
22. Wang L, Lawrence MS, Wan Y et al (2011) SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med 26:2497–2506
23. Fabbri G, Rasi S, Rossi D et al (2011) Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. J Exp Med 7:1389–1401
24. Rozman C, Montserrat E (1995) Chronic lymphocytic leukemia. N Engl J Med 16:1052–1057
25. Zenz T, Mertens D, Kuppers R et al (2010) From pathogenesis to treatment of chronic lymphocytic leukaemia. Nat Rev Cancer 1:37–50
26. Hamblin TJ, Davis Z, Gardiner A et al (1999) Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 6:1848–1854
27. Kan Z, Jaiswal BS, Stinson J et al (2010) Diverse somatic mutation patterns and pathway alterations in human cancers. Nature 7308:869–873
28. Grabher C, von Boehmer H, Look AT (2006) Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia. Nat Rev Cancer 5:347–359
29. Weng AP, Ferrando AA, Lee W et al (2004) Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science 5694:269–271
30. Palomero T, Lim WK, Odom DT et al (2006) NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc Natl Acad Sci USA 48:18261–18266
31. Tsimberidou AM, Keating MJ (2006) Richter's transformation in chronic lymphocytic leukemia. Semin Oncol 2:250–256
32. Kanai M, Hanashiro K, Kim SH et al (2007) Inhibition of Crm1-p53 interaction and nuclear export of p53 by poly(ADP-ribosyl)ation. Nat Cell Biol 10:1175–1183
33. Wang W, Budhu A, Forgues M et al (2005) Temporal and spatial control of nucleophosmin by the Ran–Crm1 complex in centrosome duplication. Nat Cell Biol 8:823–830
34. Ranganathan P, Yu X, Na C et al (2012) Pre-clinical activity of a novel CRM1 inhibitor in acute myeloid leukemia. Blood. doi:10.1182/blood-2012-04-423160
35. O'Neill LA, Bowie AG (2007) The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. Nat Rev Immunol 5:353–364
36. Ngo VN, Young RM, Schmitz R et al (2011) Oncogenically active MYD88 mutations in human lymphoma. Nature 7332:115–119
37. Kroll J, Shi X, Caprioli A et al (2005) The BTB-kelch protein KLHL6 is involved in B-lymphocyte antigen receptor signaling and germinal center formation. Mol Cell Biol 19:8531–8540
38. Dohner H, Stilgenbauer S, Benner A et al (2000) Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 26:1910–1916
39. Klein U, Lia M, Crespo M et al (2010) The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell 1:28–40
40. Calin GA, Dumitru CD, Shimizu M et al (2002) Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 24:15524–15529
41. Golas MM, Sander B, Will CL et al (2003) Molecular architecture of the multiprotein splicing factor SF3b. Science 5621:980–984
42. Wu L, Multani AS, He H et al (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. Cell 1:49–62
43. Hockemeyer D, Daniels JP, Takai H et al (2006) Recent expansion of the telomeric complex in rodents: two distinct POT1 proteins protect mouse telomeres. Cell 1:63–77
44. Nagarajan P, Onami TM, Rajagopalan S et al (2009) Role of chromodomain helicase DNA-binding protein 2 in DNA damage response signaling and tumorigenesis. Oncogene 8:1053–1062
45. Prazeres H, Torres J, Rodrigues F et al (2011) Chromosomal, epigenetic and microRNA-mediated inactivation of LRP1B, a modulator of the extracellular environment of thyroid cancer cells. Oncogene 11:1302–1317
46. Quesada V, Ramsay AJ, Lopez-Otin C (2012) Chronic lymphocytic leukemia with SF3B1 mutation. N Engl J Med 26:2530
47. Rossi D, Bruscaggin A, Spina V et al (2011) Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood 26:6904–6908
48. Yoshida K, Sanada M, Shiraishi Y et al (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 7367:64–69
49. Graubert TA, Shen D, Ding L et al (2012) Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet 1:53–57
50. Papaemmanuil E, Cazzola M, Boultwood J et al (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med 15:1384–1395
51. David CJ, Manley JL (2010) Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. Genes Dev 21:2343–2364
52. Folco EG, Coil KE, Reed R (2011) The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. Genes Dev 5:440–444
53. Corrionero A, Minana B, Valcarcel J (2011) Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. Genes Dev 5:445–459
54. Chen AA, Marsit CJ, Christensen BC et al (2009) Genetic variation in the vitamin C transporter, SLC23A2, modifies the risk of HPV16-associated head and neck cancer. Carcinogenesis 6:977–981
55. Bulwin GC, Heinemann T, Bugge V et al (2006) TIRC7 inhibits T cell proliferation by modulation of CTLA-4 expression. J Immunol 10:6833–6841
56. Yu B, Zhou X, Li B et al (2011) FOXP1 expression and its clinicopathologic significance in nodal and extranodal diffuse large B-cell lymphoma. Ann Hematol 6:701–708
57. Brown PJ, Ashe SL, Leich E et al (2008) Potentially oncogenic B-cell activation-induced smaller isoforms of FOXP1 are highly expressed in the activated B cell-like subtype of DLBCL. Blood 5:2816–2824
58. Garbelli A, Radi M, Falchi F et al (2011) Targeting the human DEAD-box polypeptide 3 (DDX3) RNA helicase as a novel strategy to inhibit viral replication. Curr Med Chem 20:3015–3027
59. Yedavalli VS, Neuveut C, Chi YH et al (2004) Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell 3:381–392
60. van der Maarel SM, Scholten IH, Huber I et al (1996) Cloning and characterization of DXS6673E, a candidate gene for X-linked mental retardation in Xq13.1. Hum Mol Genet 7:887–897
61. Xiao S, Nalabolu SR, Aster JC et al (1998) FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukaemia/lymphoma syndrome. Nat Genet 1:84–87
62. Grenard P, Bates MK, Aeschlimann D (2001) Evolution of transglutaminase genes: identification of a transglutaminase gene cluster on human chromosome 15q15. Structure of the gene encoding transglutaminase X and a novel gene family member, transglutaminase Z. J Biol Chem 35:33066–33078
63. Stratton MR (2011) Exploring the genomes of cancer cells: progress and promise. Science 6024:1553–1558
64. Service RF (2006) Gene sequencing. The race for the $1000 genome. Science 5767:1544–1546
65. Maheswaran S, Sequist LV, Nagrath S et al (2008) Detection of mutations in EGFR in circulating lung-cancer cells. N Engl J Med 4:366–377
66. Leary RJ, Kinde I, Diehl F et al (2010) Development of personalized tumor biomarkers using massively parallel sequencing. Sci Transl Med 20:20ra14