

The official journal of

INTERNATIONAL FEDERATION OF PIGMENT CELL SOCIETIES · SOCIETY FOR MELANOMA RESEARCH

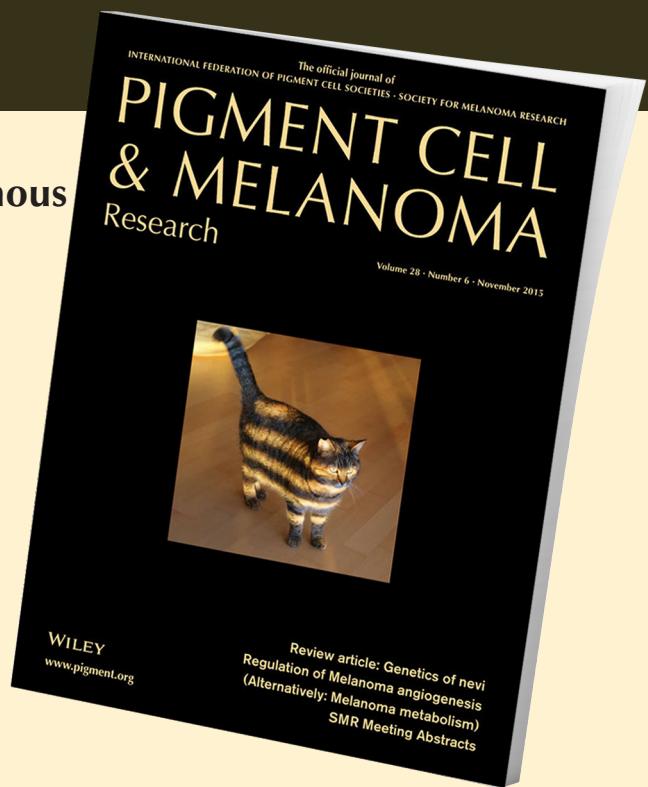
PIGMENT CELL & MELANOMA Research

The functional relevance of somatic synonymous mutations in melanoma and other cancers

Valer Gotea, Jared J. Gartner, Nouar Qutob, Laura Elnitski
and Yardena Samuels

DOI: 10.1111/pcmr.12413

Volume 28, Issue 6, Pages 673–684



If you wish to order reprints of this article,
please see the guidelines [here](#)

EMAIL ALERTS

Receive free email alerts and stay up-to-date on what is published
in Pigment Cell & Melanoma Research – [click here](#)

Submit your next paper to PCMR online at <http://mc.manuscriptcentral.com/pcmr>

Subscribe to PCMR and stay up-to-date with the only journal committed to publishing
basic research in melanoma and pigment cell biology

As a member of the IFPCS or the SMR you automatically get online access to PCMR. Sign up as
a member today at www.ifpcs.org or at www.societymelanomaresearch.org

To take out a personal subscription, please [click here](#)

More information about Pigment Cell & Melanoma Research at www.pigment.org

The functional relevance of somatic synonymous mutations in melanoma and other cancers

Valer Gotea^{1,*}, Jared J. Gartner^{2,*}, Nouar Qutob^{3,*}, Laura Elnitski¹ and Yardena Samuels³

1 Translational and Functional Genomics Branch, National Human Genome Research Institute, NIH, Bethesda, MD, USA **2** Surgery Branch, National Cancer Institute, NIH, Bethesda, MD, USA **3** Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel

CORRESPONDENCE Y. Samuels and L. Elnitski e-mails: Yardena.samuels@weizmann.ac.il and elnitski@mail.nih.gov

*These authors contributed equally to this work

KEYWORDS synonymous mutations/cancer/melanoma/gene splicing/miRNA regulation/secondary structure/protein translation/protein folding

PUBLICATION DATA Received 4 April 2015, revised and accepted for publication 19 August 2015, published online 24 August 2015

doi: 10.1111/pcmr.12413

Summary

Recent technological advances in sequencing have flooded the field of cancer research with knowledge about somatic mutations for many different cancer types. Most cancer genomics studies focus on mutations that alter the amino acid sequence, ignoring the potential impact of synonymous mutations. However, accumulating experimental evidence has demonstrated clear consequences for gene function, leading to a widespread recognition of the functional role of synonymous mutations and their causal connection to various diseases. Here, we review the evidence supporting the direct impact of synonymous mutations on gene function via gene splicing; mRNA stability, folding, and translation; protein folding; and miRNA-based regulation of expression. These results highlight the functional contribution of synonymous mutations to oncogenesis and the need to further investigate their detection and prioritization for experimental assessment.

The identification of cancer-causing mutations and the corresponding functionally impacted processes represents the main goal of cancer genomics. The inception of large collaborative efforts such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) has led to the discovery of numerous causal or driver mutations in many cancer types. Nevertheless, tumors continue to be found in the absence of conspicuous mutational events, such as nucleotide substitutions, translocations, or copy number variants involving genes with well-established tumorigenic connections. The lack of clear driver events in some cancers motivates the search for somatically acquired events that are more rare or have less obvious functional consequences but mechanistically converge on genes and pathways involved in oncogenesis and tumor progression.

Always in focus: non-synonymous mutations

The identification of the molecular basis of human cancers (Bishop, 1987; Poiesz et al., 1980; Stehelin et al., 1976; Varmus, 1984) has fueled the search for cancer-related genes and corresponding mutational events with

functional oncogenic relevance. Although identification of such genes in the past has relied on low-throughput techniques, such as linkage mapping (Bronner et al., 1994; Fishel et al., 1993; Leach et al., 1993; Peltomaki et al., 1993), the genomic era has enabled the search for somatically acquired mutations in cancers, and consequently for cancer-related genes, by directly comparing the sequence of cancer genomes with a reference genome sequence (Futreal et al., 2001). Genes linked to oncogenesis can be functionally grouped into two broad categories, oncogenes and tumor suppressors (Vogelstein et al., 2013). Oncogenes initiate and accelerate the tumorigenic process in the context of gain-of-function mutations (i.e., those leading to increased expression), whereas tumor suppressor genes confer growth advantage to cells upon acquiring loss-of-function mutations.

Recent advances in sequencing technology have enabled the discovery of an increasingly large number of functionally important somatic mutations in many cancer types (Berger et al., 2012; Davies et al., 2002; Greenman et al., 2007; Hodis et al., 2012; Kandoth et al., 2013; Krauthammer et al., 2012; Nikolaev et al., 2012; Parsons et al., 2008; Pleasance et al., 2010; Samuels et al., 2004; Stark et al., 2012; Wei et al., 2011). In turn,

these mutations have allowed the discovery of novel cancer-related genes, such as *BRAF* (Davies et al., 2002), *PIK3CA* (Samuels et al., 2004), and *IDH1* (Parsons et al., 2008). They have also revealed the dual functional nature of some genes, such as *NOTCH1*, which acts as an oncogene in T-cell acute lymphoblastic leukemia (Weng et al., 2004), and as a tumor suppressor gene in squamous cell carcinoma (Agrawal et al., 2011; Stransky et al., 2011). Additionally, these studies have reinforced the non-random nature of mutational patterns in cancer-related genes. Specifically, driver mutations in oncogenes tend to be recurrent (i.e., found in multiple samples) and to occur in hot spots that correspond to specific protein functional domains. One of the best-known examples of oncogenic hot spots is the codon 600 of *BRAF* (Davies et al., 2002), where most frequent mutations change the encoded amino acid from valine to glutamic acid (V600E). These mutations are common in many cancer types, including melanoma and colorectal and ovarian cancers. Other well-known recurrent mutations include NRAS mutations, such as Q61R and Q61K, which are frequently found in melanoma (Curtin et al., 2005; Lee et al., 2011; Platz et al., 2008); KRAS G12D, which is the most common KRAS mutation in colon and pancreatic cancers (Kim et al., 2011; Neumann et al., 2009); and IDH1 R132H, together with other mutations of residue R132, in gliomas (Yan et al., 2009). In the case of tumor suppressor genes, inactivating mutations occur throughout the gene without preference for mutational hot spots. Such mutations can affect tumor suppressor genes in several ways, including premature truncation by nonsense mutations, alteration of function through the accumulation of missense mutations, and removal or truncation of important functional regions by insertions or deletions (Kamb et al., 1994; Lopez et al., 2012; Wei et al., 2011). It is important to note that all these mutations have an obvious impact on the protein product through either amino acid replacement or protein truncation. This observation also extends to the mutations considered in the '20/20' test designed to classify genes into oncogenes or tumor suppressors (Vogelstein et al., 2013). It follows that most cancer studies continue to focus on finding frequently recurring missense, nonsense, and insertion/deletion mutations with obvious protein impact; more rare occurrences of other types of driver mutations are overlooked as a consequence.

A case for a role of synonymous mutations in cancer?

One particular class of mutations has been consistently overlooked in cancer studies, namely synonymous mutations or substitutions. They consist of single nucleotide changes in gene coding sequences that do not affect the amino acid encoded by the affected codon. For this reason, they are usually referred to as either synonymous or silent mutations. Many cancer studies completely

ignore synonymous mutations (e.g., Agrawal et al., 2012; Sankin et al., 2014), whereas others use them to build neutral background models of mutation clustering for the detection of activating mutations (Tamborero et al., 2013). Using an individual as a point of reference, such mutations can be defined as somatic, or de novo, mutations when acquired during one's lifetime, or germline mutations when inherited from either parent. In the latter case, they represent a subset of single nucleotide polymorphisms (SNPs) segregating in the population, and are sometimes referred to as sSNPs.

Apart from their mute effect on the encoded protein, the disregard for such DNA modifications may be rooted in the assumption that synonymous sites are not subject to selection (Kimura, 1977; King and Jukes, 1969). Accumulating evidence of selection acting on synonymous codons has challenged the assumption of neutrality for synonymous sites (Iida and Akashi, 2000; Shields et al., 1988), although it has been argued that in mammals it is more difficult to distinguish patterns of neutral evolution from signatures of selective pressure due to traits with minor phenotypic effects (Duret, 2002). Various studies have also shown that the rate of evolution at synonymous sites is lower than the rate of evolution observed at cognate pseudogenes or intergenic sequences (Bustamante et al., 2002; Hellmann et al., 2003). Specific examples have provided molecular mechanisms, such as gene splicing, that could explain this increased evolutionary pressure (Hurst and Pal, 2001; Orban and Olah, 2001; Pagani et al., 2005; Parmley et al., 2006). Some studies estimate that up to 40% of synonymous mutations are subject to purifying selection (Hellmann et al., 2003). Based on comparisons between human and chimp, researchers have estimated that 90% of synonymous mutations are deleterious, albeit with weak fitness effects (Lu and Wu, 2005). These results agree with the weak signatures of selection detected at synonymous sites in human population studies (Comeron, 2006). More recent comparisons with murids suggest that approximately 20% of mutations at synonymous sites are effectively selected against in hominids (Eory et al., 2010; Keightley et al., 2011). It has become clear that synonymous mutations are not entirely neutral genetic passengers, and therefore, their phenotypic contribution needs to be reevaluated. But what are the molecular underpinnings of their functional importance?

Functional impact of synonymous mutations

Evidence for functional molecular consequences of synonymous mutations has been accumulating for over three decades, with discoveries invariably prompted by the link between synonymous mutations and various disease phenotypes. Several detailed reviews of molecular mechanisms involving synonymous mutations have been published (Bali and Bebok, 2015; Cartegni et al., 2002;

Chamary et al., 2006; Hunt et al., 2014; Parmley and Hurst, 2007; Plotkin and Kudla, 2011; Sauna and Kimchi-Sarfaty, 2011, 2013; Sauna et al., 2007; Shabalina et al., 2013), highlighting their often overlooked importance for human health. Mechanisms of interference with gene function appear to be diverse, ranging from impacting mRNA splicing to protein translation (Figure 1). Here, we provide a brief overview of these mechanisms.

Splicing regulation is the mechanism for which the largest set of evidence exists. The impact of mutations (including synonymous ones) on splicing has long been known (Cartegni et al., 2002; Krawczak et al., 1992; Treisman et al., 1983). The mutations with the most obvious impact are those located next to splice junctions (Krawczak et al., 1992), and examples of such synonymous mutations are linked to various syndromes, including acute intermittent porphyria (Grandchamp et al., 1989), Tay–Sachs disease (Akli et al., 1990), phenylketonuria (Chao et al., 2001), and von Hippel–Lindau disease (Martella et al., 2006). However, some of the earliest examples of synonymous mutations linked to aberrant splicing came from the discovery of mutations in the β -globin gene in β -thalassemia patients. Among the several variants found to impact the function of the human β -globin gene, one was a synonymous substitution in codon 24 (Goldsmith et al., 1983). This substitution activated an exonic donor splice site located in codon 25, which led to truncation of the exon by 16 nucleotides.

This mutation is associated with a 75% decrease in normally processed β -globin mRNA, which could be explained by mRNA sequences being subject to nonsense-mediated decay due to truncation-induced frameshift. Activation of exonic cryptic or de novo splice sites by synonymous mutations is a well-defined molecular mechanism that has been linked to several other syndromes as well. For example, an adenine-to-guanine transition in the growth hormone receptor gene (*GHR*) activates an exonic donor splice site that results in an in-frame deletion of 24 nucleotides from exon 6 (Berg et al., 1992). Although the deletion does not change the reading frame for the rest of the protein, it does eliminate eight amino acids from the extracellular domain of GHR, conferring resistance to growth hormone, a condition known as Laron syndrome. Other conditions caused by the activation of alternative splice sites include Crouzon syndrome (Del Gatto and Breathnach, 1995; Li et al., 1995) and Rett syndrome (Sheikh et al., 2013). We have documented synonymous mutations with the same effect in the *CFTR* gene (Scott et al., 2012).

One less obvious mechanism appeared to be in play in the case of mutations located far from exon junctions and which do not activate cryptic or de novo splice sites. Instead, they cause the skipping of entire exons, as has been reported in cases of acute intermittent porphyria (Llewellyn et al., 1996) and Marfan syndrome (Liu et al., 1997). The opposite phenomenon, namely excess exon

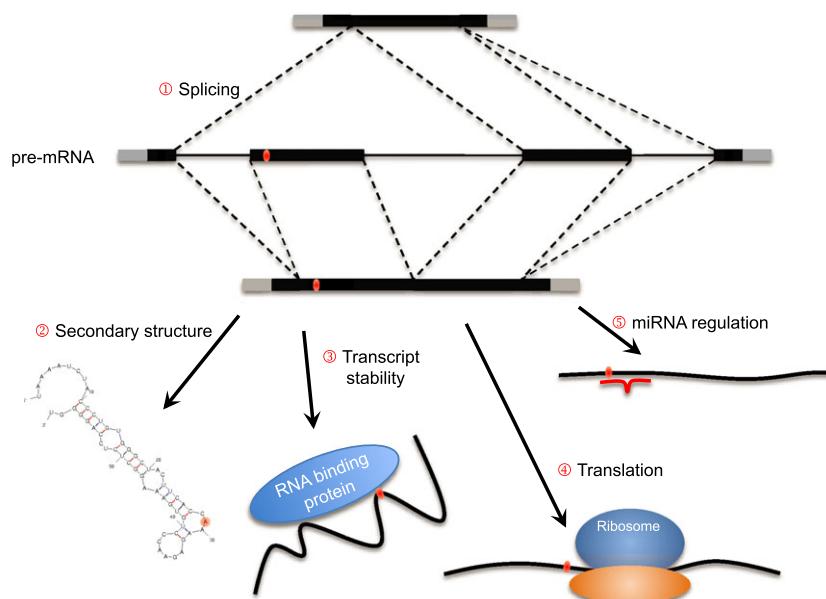


Figure 1. Conceptual representation of molecular mechanisms affected by synonymous mutations. The orange oval denotes a synonymous mutation. (1) Interference with splicing could result in multiple outcomes, such as exon skipping, exon truncation, and retention of exons at higher rates. Exon skipping is illustrated here as the outcome with the most obvious impact on the mature mRNA sequence. (2) Synonymous mutations can affect the folded secondary structure of the mature transcript resulting in altered transcript stability. (3) Stability of the transcript can also be affected by synonymous mutations through their interference with RNA-binding proteins. (4) Synonymous mutations alter the codon, which can result in either increased or decreased translation rates depending on the relative abundance of the corresponding tRNA molecules in the cell; ultimately, cotranslational folding may be affected, resulting in misfolded proteins. (5) Synonymous mutations can add or subtract miRNA binding sites, leading to altered levels of expression.

inclusion, has been observed in the case of one synonymous mutation in a patient suffering from dementia with parkinsonism, wherein disruption of a splicing silencer was suspected (D'souza et al., 1999). Other examples of synonymous mutations disrupting exonic splicing regulatory sequences (i.e., splicing silencers and enhancers; ESRs) include cases of spinal muscular atrophy (Cartegni and Krainer, 2002), cystic fibrosis (Pagani et al., 2005), and Treacher Collins syndrome (Macaya et al., 2009). Historically, the functional importance of ESRs was highlighted by the discovery of signatures of purifying selection at synonymous sites in *BRCA1* (Hurst and Pal, 2001), which were hypothesized to be due to the presence of codon-embedded splicing regulatory elements (Orban and Olah, 2001). Subsequent larger studies have provided further evidence in support of this hypothesis (Carlini and Genut, 2006; Chamary et al., 2006; Parmley et al., 2006), and recent figures estimate that 4% of synonymous mutations are deleterious owing to their disruption of ESRs, with direct implications for aberrant gene splicing (Caceres and Hurst, 2013).

The secondary structure of the transcript is another level of regulation impacted by synonymous mutations (Chamary and Hurst, 2005). Shen et al. (1999) have shown that synonymous mutations in two different genes, *AARS* and *RPA1*, lead to different folding patterns of the processed mRNA. Similarly, synonymous mutations in the human dopamine receptor D2 (*DRD2*) alter the predicted folding of the transcript, leading to a decrease in mRNA stability and translation and ultimately to decreased dopamine-induced *DRD2* expression (Duan et al., 2003). These mutations in *DRD2* are associated with schizophrenia and alcoholism. The same study highlights the compensatory role of some synonymous mutations that have an indirect functional effect by canceling the detrimental effect of other mutations.

Additional studies illustrate converging effects of synonymous mutations at the level of protein through antagonistic effects at the level of mRNA. Specifically, reduction in protein production could also result from increased stability of the mRNA secondary structure. For example, a synonymous mutation in *COMT*, a key regulator of pain perception, cognitive function, and affective mood, is associated with the most stable secondary mRNA structure, the lowest amount of translated protein, and the lowest enzymatic activity (Nackley et al., 2006). Site-directed mutagenesis has shown that mutations that reduce the stability of the secondary structure result in increases in *COMT* levels and enzymatic activity. However, other mechanisms that do not involve the mRNA secondary structure, such as differential microRNA (miRNA) targeting or interaction with ribonucleoprotein (RNP) complexes, could also explain this result and need to be ruled out to fully accept the hypothesis of increased stability of mRNA secondary structure.

Another example of a mutation affecting the mRNA secondary structure is provided by the most common

mutation associated with cystic fibrosis. This is a 3-bp deletion in the cystic fibrosis transmembrane conductance regulator (CFTR) gene and is usually referred to as ΔF508. This mutation not only eliminates residue 508, but it leads to the replacement of the last nucleotide of codon 507 (isoleucine) by the last nucleotide of codon 508, which is a synonymous change (codon ATC is replaced by ATT). This change drastically alters the mRNA folding structure, resulting in a translational pause, which consequently leads to protein misfolding and degradation (Bartoszewski et al., 2010). In bacteria, codons that prevent secondary mRNA structures from forming are favored at gene starts (Bentle et al., 2013; Kudla et al., 2009). A similar effect has been observed in human as well (Li and Qu, 2013), although it remains to be determined whether the bias favoring certain nucleotide variants at gene starts is due to their effect on mRNA secondary structures.

Interaction with RNA-binding proteins can also be influenced by synonymous mutations through affecting protein-specific interaction sites, ultimately altering the stability of mRNA species. Capon et al. (2004) have shown that a synonymous mutation in corneodesmosin (*CDSM*) increases mRNA stability. The synonymous mutation decreases affinity for a 39-kDa cytoplasmic RNA-binding protein. The effect of decreased binding is increased stability of the *CDSM* transcript. The increased stability of the transcript cannot be attributed to altered mRNA folding, as the predicted secondary structures of the reference and variant transcripts are virtually identical (Figure 2). The authors conclude that this synonymous mutation increases susceptibility to psoriasis. An antagonistic example, in which synonymous mutations reduce transcript stability, is provided by synonymous mutations that decrease the ability of *SOD1* mRNA to form RNP complexes (Ge et al., 2006). Mutations in *SOD1* account for approximately 15% of inherited mutations that cause amyotrophic lateral sclerosis. *SOD1* mRNA forms neuronal tissue-specific RNP complexes, and mutations impeding the formation of such complexes decrease mRNA stability and, consequently, mRNA levels. However, the specific proteins that interact with the *SOD1* mRNA remain to be identified.

Protein translation and folding are influenced by synonymous mutations by changing the synonymous codons. In turn, they determine the corresponding tRNA molecules, whose abundance has been linked to translation efficiency (Bulmer, 1991; Dos Reis et al., 2004; Gingold and Pilpel, 2011; Gustafsson et al., 2004; Stoletzki and Eyre-Walker, 2007; Waldman et al., 2011). Spencer et al. (2012) have illustrated this phenomenon in bacteria, in which a luciferase expression construct engineered with synonymous mutations to achieve codon optimality increases translation rates. An unexpected consequence of the increase in translation rate is reduced luciferase activity. A reengineered product with synonymous mutations that mimic the codon usage of

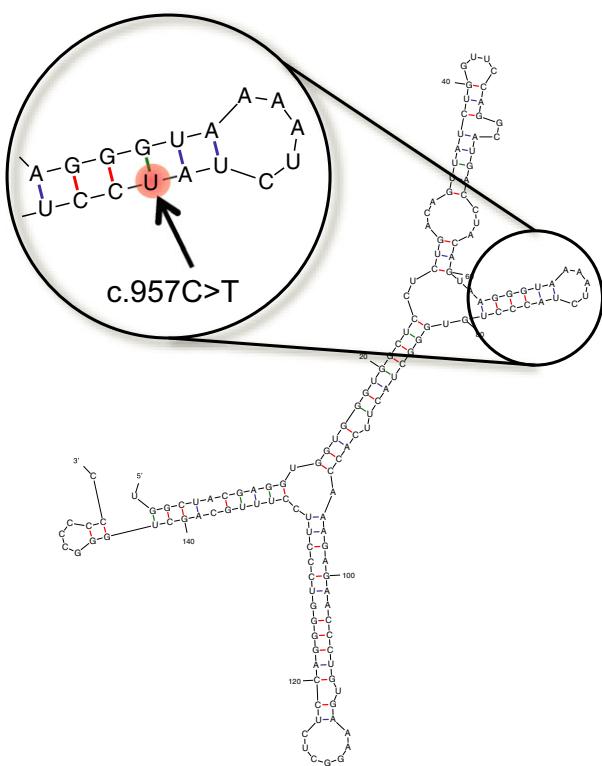


Figure 2. Partial view of the secondary structure of the *CDSN* transcript. The structure was predicted using mfold (<http://mfold.rna.albany.edu/?q=mfold>) and 151 bps (corresponding to coordinates chr6:31084360-31084510 in the hg19 assembly of the human genome) centered on the position of the synonymous mutation c.957C>T (rs1062470). The structure corresponding to the synonymous mutation is predicted to be identical to the wild-type structure (the mutated position is highlighted in the enlarged inset), with the exception of the level of free energy, which is slightly lower for the structure corresponding to the reference C allele ($\Delta G = -46.80$ kcal/mol) relative to the structure corresponding to the synonymous T variant ($\Delta G = -44.20$ kcal/mol).

Drosophila melanogaster (the closest relative to the firefly) results in a luciferase twice as active as the wild-type version, supporting a mechanism of protein folding that is reliant upon ‘ribosomal rhythm’, including pauses. These results agree with the codon optimality selection for protein folding that has been documented in eukaryotes (O’brien et al., 2012; Pechmann and Frydman, 2013; Tsai et al., 2008), although the relationship between tRNA abundance and translation speed remains contentious (Charneski and Hurst, 2013; Ingolia et al., 2011).

Skewed translation rates are proposed to alter the function of P-glycoprotein (P-gp) in human. P-gp is the protein product of *ABCB1* [the ATP-binding cassette, subfamily B (MDR/TAP), member 1 gene], also known as *MDR1* (Kimchi-Sarfaty et al., 2007). In this case, a synonymous mutation does not affect mRNA and protein levels but introduces a rare codon that alters the cotranslational folding and insertion of P-gp into the membrane. As a result, P-gp’s interactions with drugs and inhibitors such as cyclosporin A and verapamil are

altered. Likewise, translation efficiency influenced by synonymous mutations has been linked to other conditions, such as phenylketonuria (through a synonymous mutation in *PAH*) and Best’s macular dystrophy (through a synonymous mutation in *BEST1*) (Waldman et al., 2011), as well as lack of response to Herceptin treatment of patients with HER2-positive breast cancer (Griseri et al., 2011).

Affinity of miRNA binding sites is also influenced by synonymous mutations, which can change the nucleotide sequence recognized by miRNAs. The possibility that synonymous mutations interfere with miRNA binding was first proposed in 2006 (Chamary et al., 2006). The first theoretical evidence supporting this hypothesis followed soon afterward, when Hurst (2006) showed that gene sequences that are part of putative miRNA pairing domains exhibit a significantly lower rate of evolution at synonymous sites compared to the rest of gene sequences. The case of the synonymous variant rs10065172 in *IRGM* is illustrative of the dismissal of functional relevance of synonymous variants. Although this variant is strongly associated with Crohn’s disease in individuals of European descent, its functional implications were disregarded as non-causative in favor of a 20-kb deletion upstream of *IRGM* that occurred in perfect disequilibrium with the synonymous variant (Mccarroll et al., 2008). Later, an alternative hypothesis was proposed for the role of this mutation, implicating interference with posttranscriptional regulation through miRNA binding. This hypothesis was based on computational predictions that suggested the disruption of miR-196A and miR-196B binding sites (Brest et al., 2011). Experimental evidence showed that these miRNAs were overexpressed in the inflammatory intestinal epithelia of Crohn’s disease patients, resulting in the downregulation of the *IRGM* variant carrying the protective C allele but not of the variant carrying the T allele. As a result, intracellular replication of Crohn’s disease-associated *E. coli* is less effectively controlled, suggesting direct causal involvement by the synonymous variant.

Synonymous mutations in cancer

The many examples of synonymous mutations implicated in various human diseases, including heritable conditions (Sauna and Kimchi-Sarfaty, 2011), suggest parallel functional effects in cancers. Awareness that somatic synonymous mutations may include causal variants, that is, driver mutations, is beginning to take hold. Two recent studies provide evidence for the causal involvement of synonymous mutations in melanoma (Gartner et al., 2013), as well as other cancer types (Supek et al., 2014), which should provide added incentive to further investigate of synonymous mutations as driver events in cancers.

A recent study of 29 melanoma exomes and genomes revealed 16 recurrent somatic synonymous mutations (Gartner et al., 2013). To determine the prevalence of

these recurrent synonymous mutations, the authors extended the panel to include over 150 additional melanoma samples. The study revealed two highly recurrent mutations, one in *OR4C3* and one in *BCL2L12*. Both recurred at a rate greater than expected by chance, implying that these mutations underwent selection during tumor development. Given that *BCL2L12* has previously been implicated in tumor development, the F17F mutation in *BCL2L12* was further investigated in another expanded panel of tumor samples. Of a total of 256 screened samples, 10 contained the mutation. Sequenom MALDI-TOF was used to determine the abundance of transcripts for each allele using a minimum 10% threshold of the mutant allele peak, and showed that the variant allele was more abundantly expressed than the wild-type allele. To test whether this increase in transcript levels increased protein expression, the authors constructed both wild-type and mutated versions of *BCL2L12* and transiently transfected them into a 293T cell line, wherein the mutated version increased both transcript and protein levels. Several miRNA target prediction programs indicated that the change in transcript expression levels was due to differential binding of hsa-miR-671-5p miRNA, which was predicted to bind the wild-type *BCL2L12* but not the mutant mRNA. Transfection of this miRNA into cell lines had no effect on the mutant *BCL2L12* RNA but significantly reduced transcript levels of the wild-type mRNA. As *BCL2L12* binds TP53 and inhibits apoptosis in glioblastoma, the authors sought to determine its functional contribution to melanoma. Both mutant and wild-type *BCL2L12* bound TP53; however, mutant cell lines expressing the synonymous mutation lost the regulatory effect of hsa-miR-671-5p miRNA, rendering them unable to properly regulate the expression of TP53-dependent target genes. The F17F mutation in *BCL2L12*, which changes the phenylalanine-encoding codon from TTC to TTT, represented an early example of a somatic synonymous mutation with a demonstrated functional role in cancer, confirming the predictions that variants can lead to cancer by interfering with miRNA binding (Ziebarth et al., 2012).

Other studies, investigating additional functional pathways, have started to emerge. A recent study by Supek et al. (2014) proposed that synonymous mutations often act as drivers in human cancers and suggested that as many as 20–50% of synonymous mutations in oncogenes are under selection. In their thorough analysis of cancer exomes and genomes, the authors provided statistical proof of the deleterious effects of synonymous mutations in cancer. To search for genes enriched for mutations, they compiled a data set containing over 200 000 missense and 100 000 synonymous mutations generated from over 3000 exome samples comprising 11 different tissue types, with 200 samples per type. As expected, they were able to identify genes previously demonstrated to be enriched for missense mutations. Unexpectedly, oncogenes carrying activating non-synony-

mous mutations also showed clear signs of enrichment for synonymous mutations. The authors noted a 23–30% excess of this mutation type in their cancer exome samples when compared with a matched control gene set. They did not observe the same enrichment in known tumor suppressor genes, and investigation of the mutational load in the UTRs of these genes indicated that local mutation rates were not the reason for the elevated number of synonymous mutations. Furthermore, the authors compared the rates of synonymous mutations in oncogenes with those in neighboring genes, as well as simulating mutations in coding regions by sampling mutations from introns and UTRs. Both comparisons robustly confirmed the enrichment in synonymous mutations observed in this subset of oncogenes.

To determine the predominant effect of the synonymous mutations, Supek et al. searched for specific biases characteristic of the synonymous mutation set. Such biases might be similar to those observed for missense driver mutations, which exhibit positional biases relative to important functional domains or amino acid positions. To address this aspect, they assembled a collection of 383 synonymous mutations found in 18 oncogenes. Of these, 16 exhibited high ($>1.5x$) enrichment for synonymous mutations (enrichment levels were significant for 11 of these genes). Two additional genes, *ALK* and *NOTCH2*, exhibited slightly lower, although significant, enrichments of 1.47x and 1.46x, respectively. These mutations preferentially targeted evolutionarily conserved sites and significantly clustered in known oncogenes; these biases were not found for mutations in tumor suppressor genes. Notably, almost a third of these mutations (126 or 32.9%) were found in a larger set of 501 melanoma samples that we assembled by combining data provided by TCGA with data from four other studies (Dutton-Regester et al., 2014; Hodis et al., 2012; Krauthammer et al., 2012; Nikolaev et al., 2012). Melanoma was the cancer type providing the highest fraction of mutations, in agreement with the high mutational load in melanomas, closely followed by lung cancers, which contribute nearly a quarter of the mutations in this set.

Using this data set, Supek et al. noted the enrichment of synonymous mutations near exon boundaries, hinting at a role in exon splicing. As exonic splicing enhancers (ESEs) and exonic splicing silencers (ESSs) are evolutionarily conserved and relevant near exon boundaries (Keren et al., 2010; Parmley et al., 2006; Yeo et al., 2004), they investigated the mutations' impact on gene splicing by analyzing the effect of mutations on ESE and ESS sequences. They found that synonymous mutations preferentially created ESEs and abolished ESSs in the considered set of oncogenes. Of the ESE gains, 43% resulted in a gain of a known binding site for the splicing factor SRSF1. Similarly, one cluster of mutations resulting in ESS losses caused the loss of a known binding site for heterogeneous nuclear ribonucleoprotein (hnRNP) H2 splicing factor. Both SRSF1 and hnRNP have previously

been implicated in tumor progression, and the authors suggest that these gains and losses could alter the regulation of splicing and aid tumor development.

Based on the excess of synonymous mutations in oncogenes relative to matched gene sets, Supek et al. propose that nearly half of synonymous mutations in oncogenes could act as driver mutations. To evaluate the impact of synonymous mutations on gene splicing, they have also analyzed the relation between the presence of synonymous mutations and exon usage in the case of samples with available RNA-Seq data. For this purpose, they used 131 TCGA samples that have at least one synonymous mutation in any of the 18 oncogenes mentioned above, amounting to a total of 162 sample-gene pairs, 52 of which include melanoma samples. The association between synonymous mutations and exon usage was evaluated with the Mahalanobis outlier measure, which is in effect a multivariate version of the Z score. Of the 162 pairs, 33 (20.4%) exhibit significant Mahalanobis outlier measure (see Figure 5A and 5E in Supek et al., 2014), with melanoma not being significantly different from other cancer types in this respect (13 of 52, or 25%; $P = 0.4$, two-sided Fisher's exact test). However, the authors do not make clear whether the predominant form of splicing alteration involves increased exon skipping or exon retention.

To clarify this aspect, we analyzed the potential effect of all these mutations on splicing using the program SPANR (Xiong et al., 2015). A total of 327 mutations were considered (56 that were not evaluated by SPANR because they were located in terminal exons), 107 of which are mutations occurring in melanoma. More than half of the analyzed mutations were predicted to increase exon skipping rates (191 of 327, or 58.4%); melanoma was consistent with other cancer types in this respect (59 of 107, or 55.1%; $P = 0.405$, two-sided Fisher's exact test). The proportions remained roughly the same if we limited our analysis to mutations found in the 131 TCGA samples analyzed for differential exon usage by Supek et al. (a total of 153 mutations, of which 47 from melanoma), with 51.1% and 61.3% of mutations increasing exon skipping rates in melanoma and other cancer types, respectively. However, in the context of oncogenes, one might expect the majority of driver mutations to increase exon retention, and thus protein levels, which would be consistent with activating mutations. This is also suggested by the results of minigene splicing assays for six mutations tested by Supek et al. (none of these occurred in melanoma), as five of them resulted in increased exon retention. This bias should be readily identifiable for mutations associated with significant differential exon usage as evaluated using RNA-Seq data (e.g., FDR cutoff of 25% for the Mahalanobis outlier measure), given their significant impact on gene splicing. We found that such mutations (a total of 33; 14 occurred in melanoma) preferentially lead to exon retention, albeit not by a significant margin relative to mutations not

associated with a significant differential exon usage (all mutations: 51.5% vs. 39.2%; $P = 0.14$, one-sided Fisher's exact test; melanoma mutations: 64.3% vs. 42.4%; $P = 0.15$). We observed no significant difference between mutations associated with significant differential exon usage in melanoma and other cancer types (64.3% vs. 42.1%; $P = 0.29$, two-sided Fisher's exact test). However, the lack of a significant bias toward exon retention could be explained by the relatively low sample size of mutations included in this study, but also by the specific functions of the genes in question. For example, the mutation most significantly associated with differential exon usage is located in a 93-bp exon of *MSI2* and is predicted to promote exon skipping (SPANR $\Delta\Psi = -1.46$), not exon retention. Elimination of 31 amino acids from *MSI2* would affect one of two conserved mRNA binding domains and thus have downstream regulatory consequences given its important role in posttranscriptional regulation of gene expression in a variety of tissues (Sakakibara et al., 2001). Its oncogenic effect is usually observed in the context of acute myeloid leukemia when overexpressed (Kharas et al., 2010). The alteration of one of the mRNA binding domains might be considered a loss or reduction rather than gain of function event, which might still be compatible with oncogenic implications given that reduction of *MSI2* expression was observed in higher relative to lower grade gastric carcinomas (Emadi-Baygi et al., 2013).

To further evaluate the potential for a driver role of the mutations in melanoma, we quantified their frequency in our enlarged set of 501 melanoma samples. We found that all mutations associated with significant differential exon usage have very low recurrence (the mutation in *MSI2* was found in two samples, whereas all the others were found in a single sample). However, it is notable that four of the 11 samples in which these mutations were found also contained the most frequent *BRAF* non-synonymous mutation (it is found in more than 40% of all 501 melanoma samples). The other seven samples contained at least one high-frequency non-synonymous mutation (i.e., found in seven or more samples). The larger set of mutations associated with non-significant differential exon usage also show very low frequencies, only five of them having frequencies greater than one (three samples was the highest observed frequency). These observations suggest that instead of being the main driver mutations, mutations that affect splicing in oncogenes may facilitate tumor progression in concert with additional mutation events, in agreement with a model in which mutations that provide growth advantage to the cell are acquired gradually (Fearon and Vogelstein, 1990; Nowell, 1976; Vogelstein and Kinzler, 2004; Vogelstein et al., 2013), in effect a generalization of the two-hit model of tumor emergence originally detailed for retinoblastoma (Knudson, 1971). These mutations also raise the intriguing question of whether the impact on splicing may in fact extend beyond a handful of

oncogenes, and whether any relationship exists between mutation frequency and impact on splicing.

Supek et al. observed enrichment of synonymous mutations and their functional effects almost exclusively in oncogenes. Nevertheless, they provided examples of synonymous mutations affecting the tumor suppressor gene *TP53*. One mutation occurred at the 3' end of exon 6, which was shown, using an *in vivo* assay, to cause activation of a cryptic splice site and retention of intronic sequence that led to shift of reading frame in the coding sequence. Two additional recurrent synonymous mutations in *TP53* at the 3' ends of exons 4 and 9 showed similar effects of directly disrupting splice junctions. These observations highlight the relevance of splicing defects beyond a restricted set of oncogenes.

Tools for computational prediction of functional effects

Advances in whole-exome and whole-genome technologies have allowed us to go from a candidate gene approach to an unbiased interrogation of the genome, expediting interpretation of sequences in a data-driven manner. However, the difficulty of analyzing the very high numbers of variants revealed through modern sequencing technologies lies in properly evaluating their functional impact. To prioritize candidates for experimental validation, the scientific community needs computational tools to highlight the most likely causal synonymous mutations. Although several tools are available for the specific evaluation of non-synonymous mutations, such as SIFT (Ng and Henikoff, 2001), PolyPhen (Adzhubei et al., 2010; Ramensky et al., 2002), CHASM (Flanagan et al., 2010; Wong et al., 2011), and InVEx (Hodis et al., 2012), no tools are specifically designed to analyze synonymous mutations. General tools can be used for this specific case, however. For example, GERP, which measures the evolutionary constraint of positions (Cooper et al., 2005; Davydov et al., 2010), can be used to identify synonymous sites of evolutionary importance. Once variants at those positions are identified in patient samples, they can be tested for functional consequences.

Analysis of a mutation's impact on gene splicing benefits from the development of many tools aimed at evaluating the strength of splice sites or the impact of variants on splicing, such as MaxEntScan (Yeo and Burge, 2004), SplicePort (Dogan et al., 2007), Skippy (Woolfe et al., 2010), Spliceman (Lim and Fairbrother, 2012), MutPred Splice (Mort et al., 2014), and SPANR (Xiong et al., 2015). MaxEntScan and SplicePort measure the strength of canonical splice sites based on proximal sequence features; the effect of specific mutations can be interpreted by comparing the scores of splice sites using reference and variant alleles. Spliceman and Skippy use changes in hexamers surrounding a given variant to predict the likelihood of splicing disruption and exon skipping, respectively. Whereas Spliceman uses data

from 11 genomes to define splice site features, Skippy compares the context-specific effect of a mutation to a known set of exon skipping events. The program MutPred Splice employs a much larger set of variants (more than 1000) associated with disruption of splicing (although not necessarily skipping). SPANR takes tissue-specific function into account, predicting the change in the fraction of transcripts with a specific exon spliced in over a set of 16 tissues. It is important to keep in mind that all tools have non-negligible rates of false-positive predictions. Therefore, experimental testing should be prioritized for cases on which multiple tools converge with similar predictions.

Several other programs exist to analyze the effects of synonymous mutations on RNA transcripts. Changes in mRNA folding patterns can be predicted with mfold (Zuker, 1989, 2003), by comparing secondary structures predicted for reference and variant alleles. Various tools also exist to evaluate the presence of miRNA target sites, such as miRanda (John et al., 2004), PITA (Kertesz et al., 2007), and TargetScan (Garcia et al., 2011). A recently published tool, Silent Variant Analyzer (SilVA), provides a more integrative approach by combining data on sequence conservation, splice donor/acceptor sites, splice factor motifs, RNA-folding energy, and codon usage to rank synonymous variants with regard to their functional impact (Buske et al., 2013).

Conclusion

The influence of synonymous mutations on human disease is evident, with new examples appearing on a regular basis. Nevertheless, very little research has been conducted on the role of somatic synonymous mutations in cancers, less than a handful of studies having been published thus far. Going forward, the findings presented here should shape the way that we analyze the abundance of data being generated. Synonymous mutations hold one of the keys to uncovering novel driver mutations and potential novel oncogenes. The challenges of identifying causal synonymous mutations loom large, but *in silico* approaches can be used to prioritize candidates prior to any experimental undertaking. Refining the functional relevance of synonymous mutations should take into account the known properties of driver mutations, such as their recurrence in multiple patients, their localization at known oncogene mutational hot spots or functional sites evolving under selective pressure. Overall, the evidence presented herein should help synonymous mutations, especially recurring events, no longer evade the suspicion of functional importance, and effectively remove the 'silent' label too often attached to them.

Acknowledgements

We thank F. Supek and B. Lehner for providing relevant data from their study (Supek et al., 2014). This work was supported by the

Intramural Research Programs of the National Human Genome Research Institute (VG, LE). YS is supported by Israel Science Foundation grant numbers 1604/13 and 877/13 and by the ERC (StG-335377), the Knell Family, the Peter and Patricia Gruber Award, and Gideon Hamburger, Israel.

References

- Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249.
- Agrawal, N., Frederick, M.J., Pickering, C.R. et al. (2011). Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* 333, 1154–1157.
- Agrawal, N., Jiao, Y., Bettegowda, C. et al. (2012). Comparative genomic analysis of esophageal adenocarcinoma and squamous cell carcinoma. *Cancer Discov.* 2, 899–905.
- Akli, S., Chelly, J., Mezard, C., Gandy, S., Kahn, A., and Poenaru, L. (1990). A “G” to “A” mutation at position -1 of a 5' splice site in a late infantile form of Tay-Sachs disease. *J. Biol. Chem.* 265, 7324–7330.
- Bali, V., and Bebok, Z. (2015). Decoding mechanisms by which silent codon changes influence protein biogenesis and function. *Int. J. Biochem. Cell Biol.* 64, 58–74.
- Bartoszewski, R.A., Jablonsky, M., Bartoszewska, S., Stevenson, L., Dai, Q., Kappes, J., Collawn, J.F., and Bebok, Z. (2010). A synonymous single nucleotide polymorphism in ΔF508 *CFTR* alters the secondary structure of the mRNA and the expression of the mutant protein. *J. Biol. Chem.* 285, 28741–28748.
- Bentele, K., Saffert, P., Rauscher, R., Ignatova, Z., and Bluthgen, N. (2013). Efficient translation initiation dictates codon usage at gene start. *Mol. Syst. Biol.* 9, 675.
- Berg, M.A., Guevara-Aguirre, J., Rosenbloom, A.L., Rosenfeld, R.G., and Francke, U. (1992). Mutation creating a new splice site in the growth hormone receptor genes of 37 Ecuadorian patients with Laron syndrome. *Hum. Mutat.* 1, 24–32.
- Berger, M.F., Hodis, E., Heffernan, T.P. et al. (2012). Melanoma genome sequencing reveals frequent *PREX2* mutations. *Nature* 485, 502–506.
- Bishop, J.M. (1987). The molecular genetics of cancer. *Science* 235, 305–311.
- Brest, P., Lapaquette, P., Souidi, M. et al. (2011). A synonymous variant in *IRGM* alters a binding site for miR-196 and causes deregulation of *IRGM*-dependent xenophagy in Crohn's disease. *Nat. Genet.* 43, 242–245.
- Bronner, C.E., Baker, S.M., Morrison, P.T. et al. (1994). Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature* 368, 258–261.
- Bulmer, M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics* 129, 897–907.
- Buske, O.J., Manickaraj, A., Mital, S., Ray, P.N., and Brudno, M. (2013). Identification of deleterious synonymous variants in human genomes. *Bioinformatics* 29, 1843–1850.
- Bustamante, C.D., Nielsen, R., and Hartl, D.L. (2002). A maximum likelihood method for analyzing pseudogene evolution: implications for silent site evolution in humans and rodents. *Mol. Biol. Evol.* 19, 110–117.
- Caceres, E.F., and Hurst, L.D. (2013). The evolution, impact and properties of exonic splice enhancers. *Genome Biol.* 14, R143.
- Capon, F., Allen, M.H., Ameen, M., Burden, A.D., Tillman, D., Barker, J.N., and Trembath, R.C. (2004). A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Hum. Mol. Genet.* 13, 2361–2368.
- Carlini, D.B., and Genut, J.E. (2006). Synonymous SNPs provide evidence for selective constraint on human exonic splicing enhancers. *J. Mol. Evol.* 62, 89–98.
- Cartegni, L., and Krainer, A.R. (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in *SMN2* causes spinal muscular atrophy in the absence of *SMN1*. *Nat. Genet.* 30, 377–384.
- Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.* 3, 285–298.
- Chamary, J.V., and Hurst, L.D. (2005). Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome Biol.* 6, R75.
- Chamary, J.V., Parmley, J.L., and Hurst, L.D. (2006). Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat. Rev. Genet.* 7, 98–108.
- Chao, H.K., Hsiao, K.J., and Su, T.S. (2001). A silent mutation induces exon skipping in the phenylalanine hydroxylase gene in phenylketonuria. *Hum. Genet.* 108, 14–19.
- Charneski, C.A., and Hurst, L.D. (2013). Positively charged residues are the major determinants of ribosomal velocity. *PLoS Biol.* 11, e1001508.
- Comeron, J.M. (2006). Weak selection and recent mutational changes influence polymorphic synonymous mutations in humans. *Proc. Natl Acad. Sci. USA* 103, 6940–6945.
- Cooper, G.M., Stone, E.A., Asimenos, G., Green, E.D., Batzoglou, S., and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* 15, 901–913.
- Curtin, J.A., Fridlyand, J., Kageshita, T. et al. (2005). Distinct sets of genetic alterations in melanoma. *N. Engl. J. Med.* 353, 2135–2147.
- Davies, H., Bignell, G.R., Cox, C. et al. (2002). Mutations of the *BRAF* gene in human cancer. *Nature* 417, 949–954.
- Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* 6, e1001025.
- Del Gatto, F., and Breathnach, R. (1995). A Crouzon syndrome synonymous mutation activates a 5' splice site within the IIIc exon of the *FGFR2* gene. *Genomics* 27, 558–559.
- Dogan, R.I., Getoor, L., Wilbur, W.J., and Mount, S.M. (2007). SplicePort—An interactive splice-site analysis tool. *Nucleic Acids Res.* 35, W285–W291.
- Dos Reis, M., Savva, R., and Wernisch, L. (2004). Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res.* 32, 5036–5044.
- D'souza, I., Poorkaj, P., Hong, M., Nochlin, D., Lee, V.M., Bird, T.D., and Schellenberg, G.D. (1999). Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc. Natl Acad. Sci. USA* 96, 5598–5603.
- Duan, J., Wainwright, M.S., Comeron, J.M., Saitou, N., Sanders, A.R., Gelernter, J., and Gejman, P.V. (2003). Synonymous mutations in the human *dopamine receptor D2 (DRD2)* affect mRNA stability and synthesis of the receptor. *Hum. Mol. Genet.* 12, 205–216.
- Duret, L. (2002). Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev.* 12, 640–649.
- Dutton-Reed, K., Gartner, J.J., Emmanuel, R. et al. (2014). A highly recurrent *RPS27* 5'UTR mutation in melanoma. *Oncotarget* 5, 2912–2917.
- Emadi-Baygi, M., Nikpour, P., Mohammad-Hashem, F., Maracy, M.R., and Haghjooy-Javanmard, S. (2013). *MSI2* expression is decreased in grade II of gastric carcinoma. *Pathol. Res. Pract.* 209, 689–691.

- Eory, L., Halligan, D.L., and Keightley, P.D. (2010). Distributions of selectively constrained sites and deleterious mutation rates in the hominid and murid genomes. *Mol. Biol. Evol.* 27, 177–192.
- Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell* 61, 759–767.
- Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., and Kolodner, R. (1993). The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. *Cell* 75, 1027–1038.
- Flanagan, S.E., Patch, A.M., and Ellard, S. (2010). Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet. Test. Mol. Biomarkers* 14, 533–537.
- Futreal, P.A., Kasprzyk, A., Birney, E., Mullikin, J.C., Wooster, R., and Stratton, M.R. (2001). Cancer and genomics. *Nature* 409, 850–852.
- Garcia, D.M., Baek, D., Shin, C., Bell, G.W., Grimson, A., and Bartel, D.P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.* 18, 1139–1146.
- Gartner, J.J., Parker, S.C., Prickett, T.D. et al. (2013). Whole-genome sequencing identifies a recurrent functional synonymous mutation in melanoma. *Proc. Natl Acad. Sci. USA* 110, 13481–13486.
- Ge, W.W., Leystra-Lantz, C., Sanelli, T.R., Mclean, J., Wen, W., Strong, W., and Strong, M.J. (2006). Neuronal tissue-specific ribonucleoprotein complex formation on *SOD1* mRNA: alterations by ALS *SOD1* mutations. *Neurobiol. Dis.* 23, 342–350.
- Gingold, H., and Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. *Mol. Syst. Biol.* 7, 481.
- Goldsmith, M.E., Humphries, R.K., Ley, T., Cline, A., Kantor, J.A., and Nienhuis, A.W. (1983). "Silent" nucleotide substitution in a β -thalassemia globin gene activates splice site in coding sequence RNA. *Proc. Natl Acad. Sci. USA* 80, 2318–2322.
- Grandchamp, B., Picat, C., De Rooij, F., Beaumont, C., Wilson, P., Deybach, J.C., and Nordmann, Y. (1989). A point mutation G → A in exon 12 of the porphobilinogen deaminase gene results in exon skipping and is responsible for acute intermittent porphyria. *Nucleic Acids Res.* 17, 6637–6649.
- Greenman, C., Stephens, P., Smith, R. et al. (2007). Patterns of somatic mutation in human cancer genomes. *Nature* 446, 153–158.
- Griseri, P., Bourcier, C., Hieblot, C., Essafi-Benkhadir, K., Chamorey, E., Touriol, C., and Pages, G. (2011). A synonymous polymorphism of the Tristetraprolin (TTP) gene, an AU-rich mRNA-binding protein, affects translation efficiency and response to Herceptin treatment in breast cancer patients. *Hum. Mol. Genet.* 20, 4556–4568.
- Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends Biotechnol.* 22, 346–353.
- Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B., and Paabo, S. (2003). Selection on human genes as revealed by comparisons to chimpanzee cDNA. *Genome Res.* 13, 831–837.
- Hodis, E., Watson, I.R., Kryukov, G.V. et al. (2012). A landscape of driver mutations in melanoma. *Cell* 150, 251–263.
- Hunt, R.C., Simhadri, V.L., Iandoli, M., Sauna, Z.E., and Kimchi-Sarfaty, C. (2014). Exposing synonymous mutations. *Trends Genet.* 30, 308–321.
- Hurst, L.D. (2006). Preliminary assessment of the impact of microRNA-mediated regulation on coding sequence evolution in mammals. *J. Mol. Evol.* 63, 174–182.
- Hurst, L.D., and Pal, C. (2001). Evidence for purifying selection acting on silent sites in *BRCA1*. *Trends Genet.* 17, 62–65.
- Iida, K., and Akashi, H. (2000). A test of translational selection at 'silent' sites in the human genome: base composition comparisons in alternatively spliced genes. *Gene* 261, 93–105.
- Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147, 789–802.
- John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. *PLoS Biol.* 2, e363.
- Kamb, A., Gruis, N.A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S.V., Stockert, E., Day, R.S. 3rd, Johnson, B.E., and Skolnick, M.H. (1994). A cell cycle regulator potentially involved in genesis of many tumor types. *Science* 264, 436–440.
- Kandoth, C., McLellan, M.D., Vandin, F. et al. (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502, 333–339.
- Keightley, P.D., Eory, L., Halligan, D.L., and Kirkpatrick, M. (2011). Inference of mutation parameters and selective constraint in mammalian coding sequences by approximate Bayesian computation. *Genetics* 187, 1153–1161.
- Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11, 345–355.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39, 1278–1284.
- Kharas, M.G., Lengner, C.J., Al-Shahrour, F. et al. (2010). Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. *Nat. Med.* 16, 903–908.
- Kim, S.T., Lim Do, H., Jang, K.T. et al. (2011). Impact of *KRAS* mutations on clinical outcomes in pancreatic cancer patients treated with first-line gemcitabine-based chemotherapy. *Mol. Cancer Ther.* 10, 1993–1999.
- Kimchi-Sarfaty, C., Oh, J.M., Kim, I.W., Sauna, Z.E., Calcagno, A.M., Ambudkar, S.V., and Gottesman, M.M. (2007). A "silent" polymorphism in the *MDR1* gene changes substrate specificity. *Science* 315, 525–528.
- Kimura, M. (1977). Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* 267, 275–276.
- King, J.L., and Jukes, T.H. (1969). Non-Darwinian evolution. *Science* 164, 788–798.
- Knudson, A.G. Jr (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* 68, 820–823.
- Krauthammer, M., Kong, Y., Ha, B.H. et al. (2012). Exome sequencing identifies recurrent somatic *RAC1* mutations in melanoma. *Nat. Genet.* 44, 1006–1014.
- Krawczak, M., Reiss, J., and Cooper, D.N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* 90, 41–54.
- Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324, 255–258.
- Leach, F.S., Nicolaides, N.C., Papadopoulos, N. et al. (1993). Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75, 1215–1225.
- Lee, J.H., Choi, J.W., and Kim, Y.S. (2011). Frequencies of *BRAF* and *NRAS* mutations are different in histological types and sites of origin of cutaneous melanoma: a meta-analysis. *Br. J. Dermatol.* 164, 776–784.
- Li, Q., and Qu, H.Q. (2013). Human coding synonymous single nucleotide polymorphisms at ramp regions of mRNA translation. *PLoS ONE* 8, e59706.
- Li, X., Park, W.J., Pyeritz, R.E., and Jabs, E.W. (1995). Effect on splicing of a silent *FGFR2* mutation in Crouzon syndrome. *Nat. Genet.* 9, 232–233.
- Lim, K.H., and Fairbrother, W.G. (2012). Spliceman—a computational web server that predicts sequence variations in pre-mRNA splicing. *Bioinformatics* 28, 1031–1032.
- Liu, W., Qian, C., and Francke, U. (1997). Silent mutation induces exon skipping of fibrillin-1 gene in Marfan syndrome. *Nat. Genet.* 16, 328–329.

- Llewellyn, D.H., Scobie, G.A., Urquhart, A.J., Whatley, S.D., Roberts, A.G., Harrison, P.R., and Elder, G.H. (1996). Acute intermittent porphyria caused by defective splicing of porphobilinogen deaminase RNA: a synonymous codon mutation at -22 bp from the 5' splice site causes skipping of exon 3. *J. Med. Genet.* 33, 437–438.
- Lopez, I., Oliveira, L.P., Tucci, P., Alvarez-Valin, F., Coudry, R.A., and Marin, M. (2012). Different mutation profiles associated to P53 accumulation in colorectal cancer. *Gene* 499, 81–87.
- Lu, J., and Wu, C.I. (2005). Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. *Proc. Natl Acad. Sci. USA* 102, 4063–4067.
- Macaya, D., Katsanis, S.H., Heffron, T.W., Audlin, S., Mendelsohn, N.J., Roggenbuck, J., and Cutting, G.R. (2009). A synonymous mutation in *TCOF1* causes Treacher Collins syndrome due to mis-splicing of a constitutive exon. *Am. J. Med. Genet. A* 149A, 1624–1627.
- Martella, M., Salviati, L., Casarin, A., Trevisson, E., Opocher, G., Polli, R., Gross, D., and Murgia, A. (2006). Molecular analysis of two uncharacterized sequence variants of the *VHL* gene. *J. Hum. Genet.* 51, 964–968.
- Mccarroll, S.A., Huett, A., Kuballa, P. et al. (2008). Deletion polymorphism upstream of *IRGM* associated with altered *IRGM* expression and Crohn's disease. *Nat. Genet.* 40, 1107–1112.
- Mort, M., Sterne-Weiler, T., Li, B., Ball, E.V., Cooper, D.N., Radivojac, P., Sanford, J.R., and Mooney, S.D. (2014). MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing. *Genome Biol.* 15, R19.
- Nackley, A.G., Shabalina, S.A., Tchivileva, I.E., Satterfield, K., Korchynskyi, O., Makarov, S.S., Maixner, W., and Diatchenko, L. (2006). Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* 314, 1930–1933.
- Neumann, J., Zeindl-Eberhart, E., Kirchner, T., and Jung, A. (2009). Frequency and type of *KRAS* mutations in routine diagnostic analysis of metastatic colorectal cancer. *Pathol. Res. Pract.* 205, 858–862.
- Ng, P.C., and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Res.* 11, 863–874.
- Nikolaev, S.I., Rimoldi, D., Iseli, C. et al. (2012). Exome sequencing identifies recurrent somatic *MAP2K1* and *MAP2K2* mutations in melanoma. *Nat. Genet.* 44, 133–139.
- Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science* 194, 23–28.
- O'brien, E.P., Vendruscolo, M., and Dobson, C.M. (2012). Prediction of variable translation rate effects on cotranslational protein folding. *Nat. Commun.* 3, 868.
- Orban, T.I., and Olah, E. (2001). Purifying selection on silent sites – a constraint from splicing regulation? *Trends Genet.* 17, 252–253.
- Pagani, F., Raponi, M., and Baralle, F.E. (2005). Synonymous mutations in *CFTR* exon 12 affect splicing and are not neutral in evolution. *Proc. Natl Acad. Sci. USA* 102, 6368–6372.
- Parmley, J.L., and Hurst, L.D. (2007). How do synonymous mutations affect fitness? *BioEssays* 29, 515–519.
- Parmley, J.L., Chamary, J.V., and Hurst, L.D. (2006). Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. *Mol. Biol. Evol.* 23, 301–309.
- Parsons, D.W., Jones, S., Zhang, X. et al. (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science* 321, 1807–1812.
- Pechmann, S., and Frydman, J. (2013). Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat. Struct. Mol. Biol.* 20, 237–243.
- Peltomaki, P., Aaltonen, L.A., Sistonen, P. et al. (1993). Genetic mapping of a locus predisposing to human colorectal cancer. *Science* 260, 810–812.
- Platz, A., Egyhazi, S., Ringborg, U., and Hansson, J. (2008). Human cutaneous melanoma; a review of *NRAS* and *BRAF* mutation frequencies in relation to histogenetic subclass and body site. *Mol. Oncol.* 1, 395–405.
- Pleasance, E.D., Cheetham, R.K., Stephens, P.J. et al. (2010). A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 463, 191–196.
- Plotkin, J.B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12, 32–42.
- Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D., and Gallo, R.C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl Acad. Sci. USA* 77, 7415–7419.
- Ramensky, V., Bork, P., and Sunyaev, S. (2002). Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* 30, 3894–3900.
- Sakakibara, S., Nakamura, Y., Satoh, H., and Okano, H. (2001). Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. *J. Neurosci.* 21, 8091–8107.
- Samuels, Y., Wang, Z., Bardelli, A. et al. (2004). High frequency of mutations of the *PIK3CA* gene in human cancers. *Science* 304, 554.
- Sankin, A., Hakimi, A.A., Mikkilineni, N. et al. (2014). The impact of genetic heterogeneity on biomarker development in kidney cancer assessed by multiregional sampling. *Cancer Med.* 3, 1485–1492.
- Sauna, Z.E., and Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet.* 12, 683–691.
- Sauna, Z.E., and Kimchi-Sarfaty, C. (2013). Synonymous mutations as a cause of human genetic disease. eLS, (Chichester: John Wiley & Sons, Ltd).
- Sauna, Z.E., Kimchi-Sarfaty, C., Ambudkar, S.V., and Gottesman, M.M. (2007). The sounds of silence: synonymous mutations affect function. *Pharmacogenomics* 8, 527–532.
- Scott, A., Petrykowska, H.M., Heffron, T., Gotea, V., and Elnitski, L. (2012). Functional analysis of synonymous substitutions predicted to affect splicing of the *CFTR* gene. *J. Cyst. Fibros.* 11, 511–517.
- Shabalina, S.A., Spiridonov, N.A., and Kashina, A. (2013). Sounds of silence: synonymous nucleotides as a key to biological regulation and complexity. *Nucleic Acids Res.* 41, 2073–2094.
- Sheikh, T.I., Mittal, K., Willis, M.J., and Vincent, J.B. (2013). A synonymous change, p.Gly16Gly in *MECP2* Exon 1, causes a cryptic splice event in a Rett syndrome patient. *Orphanet J Rare Dis* 8, 108.
- Shen, L.X., Basilion, J.P., and Stanton, V.P. Jr (1999). Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc. Natl Acad. Sci. USA* 96, 7871–7876.
- Shields, D.C., Sharp, P.M., Higgins, D.G., and Wright, F. (1988). "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol. Biol. Evol.* 5, 704–716.
- Spencer, P.S., Siller, E., Anderson, J.F., and Barral, J.M. (2012). Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. *J. Mol. Biol.* 422, 328–335.
- Stark, M.S., Woods, S.L., Gartside, M.G. et al. (2012). Frequent somatic mutations in *MAP3K5* and *MAP3K9* in metastatic melanoma identified by exome sequencing. *Nat. Genet.* 44, 165–169.
- Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976). DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260, 170–173.
- Stoletzki, N., and Eyre-Walker, A. (2007). Synonymous codon usage in *Escherichia coli*: selection for translational accuracy. *Mol. Biol. Evol.* 24, 374–381.

- Stransky, N., Egloff, A.M., Tward, A.D. et al. (2011). The mutational landscape of head and neck squamous cell carcinoma. *Science* 333, 1157–1160.
- Supek, F., Minana, B., Valcarcel, J., Gabaldon, T., and Lehner, B. (2014). Synonymous mutations frequently act as driver mutations in human cancers. *Cell* 156, 1324–1335.
- Tamborero, D., Gonzalez-Perez, A., and Lopez-Bigas, N. (2013). OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics* 29, 2238–2244.
- Treisman, R., Orkin, S.H., and Maniatis, T. (1983). Specific transcription and RNA splicing defects in five cloned β -thalassaemia genes. *Nature* 302, 591–596.
- Tsai, C.J., Sauna, Z.E., Kimchi-Sarfaty, C., Ambudkar, S.V., Gottesman, M.M., and Nussinov, R. (2008). Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J. Mol. Biol.* 383, 281–291.
- Varmus, H.E. (1984). The molecular genetics of cellular oncogenes. *Annu. Rev. Genet.* 18, 553–612.
- Vogelstein, B., and Kinzler, K.W. (2004). Cancer genes and the pathways they control. *Nat. Med.* 10, 789–799.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A. Jr, and Kinzler, K.W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558.
- Waldman, Y.Y., Tuller, T., Keinan, A., and Ruppin, E. (2011). Selection for translation efficiency on synonymous polymorphisms in recent human evolution. *Genome Biol Evol* 3, 749–761.
- Wei, X., Walia, V., Lin, J.C. et al. (2011). Exome sequencing identifies *GRIN2A* as frequently mutated in melanoma. *Nat. Genet.* 43, 442–446.
- Weng, A.P., Ferrando, A.A., Lee, W., Morris, J.P.T., Silverman, L.B., Sanchez-Irizarry, C., Blacklow, S.C., Look, A.T., and Aster, J.C. (2004). Activating mutations of *NOTCH1* in human T cell acute lymphoblastic leukemia. *Science* 306, 269–271.
- Wong, W.C., Kim, D., Carter, H., Diekhans, M., Ryan, M.C., and Karchin, R. (2011). CHASM and SNVBox: toolkit for detecting biologically important single nucleotide mutations in cancer. *Bioinformatics* 27, 2147–2148.
- Woolfe, A., Mullikin, J.C., and Elnitski, L. (2010). Genomic features defining exonic variants that modulate splicing. *Genome Biol.* 11, R20.
- Xiong, H.Y., Alipanahi, B., Lee, L.J. et al. (2015). RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* 347, 1254806.
- Yan, H., Parsons, D.W., Jin, G. et al. (2009). *IDH1* and *IDH2* mutations in gliomas. *N. Engl. J. Med.* 360, 765–773.
- Yeo, G., and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* 11, 377–394.
- Yeo, G., Hoon, S., Venkatesh, B., and Burge, C.B. (2004). Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc. Natl Acad. Sci. USA* 101, 15700–15705.
- Ziebarth, J.D., Bhattacharya, A., and Cui, Y. (2012). Integrative analysis of somatic mutations altering microRNA targeting in cancer genomes. *PLoS ONE* 7, e47137.
- Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science* 244, 48–52.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415.