

Open access • Journal Article • DOI:10.1038/S41397-019-0102-4

# Pharmacogenomics, biomarker network, and allele frequencies in colorectal cancer — Source link [2]

Andrés López-Cortés, César Paz-y-Miño, Santiago Guerrero, Gabriela Jaramillo-Koupermann ...+10 more authors

Institutions: Pompeu Fabra University, Heidelberg University, German Cancer Research Center, University of Chile ...+1 more institutions

Published on: 01 Feb 2020 - Pharmacogenomics Journal (Nature Publishing Group)

Topics: Pharmacogenomics and Precision medicine

Related papers:

- · Pharmacogenomics, biomarker network and allele frequencies in colorectal cancer
- Open Targets Platform: new developments and updates two years on.
- Pathogenic Germline Variants in 10,389 Adult Cancers
- Analysis of Racial/Ethnic Representation in Select Basic and Applied Cancer Research Studies.
- Gene prioritization, communality analysis, networking and metabolic integrated pathway to better understand breast
  cancer pathogenesis



#### Pharmacogenomics, biomarker network and allele frequencies in colorectal cancer

Andrés López-Cortés<sup>1,2,3\*</sup>, César Paz-y-Miño<sup>1,\*</sup>, Santiago Guerrero<sup>1,4</sup>, Gabriela Jaramillo-Koupermann<sup>5</sup>, Dámaris P. Intriago-Baldeón<sup>6,7</sup>, Jennyfer M. García-Cárdenas<sup>1</sup>, Patricia Guevara-Ramírez<sup>1</sup>, Isaac Armendáriz-Castillo<sup>1</sup>, Paola E. Leone<sup>1</sup>, Luis Abel Quiñones<sup>2,8</sup>, Juan Pablo Cayún<sup>2,8</sup>, Néstor W. Soria<sup>9</sup>

- <sup>1</sup> Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, 170527 Quito, Ecuador
- <sup>2</sup> Latin American Society of Pharmacogenomics and Personalized Medicine
- <sup>3</sup> RNASA-IMEDIR, Computer Sciences Faculty, University of Coruna, 15071 Coruna, Spain
- <sup>4</sup> Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomics Regulation (CRG), The Barcelona Institute for Science and Technology, Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain
- <sup>5</sup> Laboratorio de Biología Molecular, Subproceso de Anatomía Patológica, Hospital de Especialidades Eugenio Espejo, 170403 Quito, Ecuador
- <sup>6</sup> Division of Molecular Genome Analysis, German Cancer Research Center (DKFZ),
   69120 Heidelberg, Germany
- <sup>7</sup> Faculty of Biosciences, Heidelberg University, 69120 Heidelberg, Germany
- <sup>8</sup> Laboratory of Chemical Carcinogenesis and Pharmacogenetics, Department of Basic-Clinical Oncology, Faculty of Medicine, University of Chile, 70111 Santiago, Chile
- <sup>9</sup> Cátedra de Biotecnología, Facultad de Ciencias Químicas, Universidad Católica de Córdoba, Unidad Asociada al CONICET, 5000 Córdoba, Argentina

## \* Corresponding authors:

## Andrés López-Cortés

Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio

Espejo, Universidad UTE, Mariscal Sucre Avenue, 170129 Quito, Ecuador.

E-mail: <u>aalc84@gmail.com</u>

## César Paz-y-Miño

Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio

Espejo, Universidad UTE, Mariscal Sucre Avenue, 170129 Quito, Ecuador.

E-mail: <a href="mailto:cesar.pazymino@ute.edu.ec">cesar.pazymino@ute.edu.ec</a>

Running title: Pharmacogenomics in colorectal cancer

### ABSTRACT

Colorectal cancer (CRC) is one of the leading causes of cancer death worldwide. Over the last decades, several studies have shown that tumor-related genomic alterations predict tumor prognosis, drug response and toxicity. These observations have led to the development of a number of precision therapies based on individual genomic profiles. As part of these approaches, pharmacogenomics analyses genomic alterations that may predict an efficient therapeutic response. Studying these mutations as biomarkers for predicting drug response is of a great interest to improve precision medicine. Here we conduct a comprehensive review of the main pharmacogenomics biomarkers and genomic alterations affecting enzyme activity, transporter capacity, channels and receptors, and therefore the new advances in CRC precision medicine to select the best therapeutic strategy in populations worldwide, with a focus on Latin America.

### **INTRODUCTION**

Colorectal cancer (CRC) is one of the leading causes of cancer death worldwide<sup>1</sup>. In the last decade, numerous exciting advances have been made to treat patients even with metastatic CRC<sup>2</sup>. However, patient-tailored therapies are still needed to overcome this disease. The advance of precision medicine requires the accurate identification of mutations driving each patient's tumor<sup>3</sup>. In this regard, genetic mutations may have a great impact on disease prognosis and therapy response. Germline mutations are heritable alterations found in individuals while somatic mutations appear after an oncogenic insult within the tumoral tissue<sup>4</sup>. As part of CRC precision medicine, pharmacogenomics allows tailoring drug therapy based on these mutations<sup>5</sup>. Thus, personalized therapy not only maximizes the drug therapeutic effects but also reduces the possibility of experiencing adverse drug reactions<sup>6</sup>. In this review, we focus primarily on the current status of pharmacogenomics in CRC, its biomarkers and allele frequencies worldwide, with a focus on Latin American populations in order to improve precision medicine.

### **Colorectal cancer oncogenomics**

CRC was one of the first solid tumors to be molecularly characterized, in whose pathogenesis several signaling pathways intervene<sup>7</sup>. Vogelstein *et al.*, described the model of progressive step-wise accumulation of epigenetic events of  $CRC^{8-11}$ . This model provides information about the role of driver mutations whose objective is to give a selective advantage for tumor progression<sup>12</sup>. In addition, the accumulation of pathogenic

mutations in the transforming growth factor- $\beta$  (TGF $\beta$ ), WNT- $\beta$ -catenin, PI3K, EGFR and downstream MAPK pathways induces CRC<sup>11,13–15</sup>.

On the other hand, the development of CRC also occurs when chromosomal instability (CIN) occurs, progress due to defects in telomere stability, chromosomal segregation and mutations in TP53 gene<sup>16</sup>. The 15% of early-stage colorectal tumors present mismatch repair-deficient (MMRd) system, triggering hypermutation and microsatellite instability (MSI)<sup>14</sup>. According to Dienstmann *et al.*, the epigenetic profile of tumors with CIN present mutations in APC, KRAS, TP53, SMAD4 and PIK3CA, promoting the formation of the non-hypermutated consensus molecular subtypes (CMSs): CMS2, CMS3 and CMS4<sup>1</sup>. Whereas tumors with MSI harbor mutations in the MSH6, RNF43, ATM, TGFBR2, BRAF and PTEN genes of the hypermutated molecular subtype CMS1<sup>1</sup>.

### A consensus of molecular subtypes

Gene expression-based subtyping is widely accepted as a relevant source of disease stratification<sup>17</sup>. Nevertheless, the translational utility is hampered by divergent results that are probably related to differences in algorithms applied to sample preparation methods, gene expression platforms and racial/ethnic disparities<sup>18,19</sup>. Inspection of the published gene expression-based CRC classification revealed an absence of a clear methodological 'gold standard'<sup>8,9,16,20–23</sup>. To facilitate clinical translation, the CRC Subtyping Consortium (CRCSC) was formed to assess the core subtype patterns among existing gene expression-based CRC subtyping algorithms<sup>18,24</sup>.

In spite of heterogeneities, subtype concordance analysis readily yielded four CMSs<sup>18</sup>, being CMS1 the immune subtype, CMS2 the canonical subtype, CMS3 the metabolic subtype and CMS4 the mesenchymal subtype (Figure 1)<sup>18,25</sup>. Upon evaluation of these classification system, Calon *et al.*, discovered that their prediction power arises from genes expressed by stromal cells that associate robustly with disease relapse<sup>26</sup>. Mesenchymal stromal cells (MSC) may represent a pivotal part of stroma in CRC, but little is known about the specific interaction of MSC in CRC<sup>27</sup>.

Recognizing that transcriptomics represents the level of high-throughput molecular data that is most intimately linked to tumor phenotype and clinical behavior, it is important to characterize the CRC genomics alterations. Tumor genomes contain thousands of mutations. However, only a few of them drive tumorigenesis by affecting driver genes, which upon alteration, confers selective growth advantage to tumor cells<sup>28</sup>. Since the identification of the first somatic mutation in human bladder carcinoma cell line (HRAS G12V)<sup>29,30</sup>, the Pan-Cancer Atlas from The Cancer Genome Atlas (TCGA) have undertaken omics analyses identifying 20 CRC driver genes (ACVR2A, AMER1, APC, ARID1A, BRAF, CTNNB1, FBXW7, GNAS, KRAS, NRAS, PCBP1, PIK3CA, PTEN, SMAD4, SMAD2, SOX9, TCF7L2, TGIF1, TP53 and ZFP36L2) that are included in The Catalogue of Somatic Mutations in Cancer (COSMIC) Cancer Gene Census (CGC) and the Cancer Genome Interpreter (CGI)<sup>31–35</sup>. CGI identifies 71 biomarkers among biallelic markers, copy number alterations (CNAs), somatic mutations, fusion genes and amplifications<sup>34</sup>. Likewise, CGI annotates CRC tumor variants that constitute state-of-art biomarkers of drug response as shown in Supplementary Table 1.

### Drugs, biomarkers and allele frequencies

According to the National Comprehensive Cancer Network (NCCN) guidelines v1.2018 and the European Society for Medical Oncology (ESMO) guidelines<sup>36–38</sup>, the two main drug categories in CRC treatment are cytotoxic and biological therapies. Cytotoxic agents are platinum derivatives (oxaliplatin), antimetabolites (5-fluorouracil and capecitabine) and antitopoisomerases (irinotecan). Biological therapy includes drugs against the epidermal growth factor receptor (EGFR) (cetuximab, panitumumab) and antiangiogenics (bevacizumab, ziv-aflibercept, ramucirumab). In addition, the recommendation includes PD-1 and PD-L1 inhibitors (nivolumumab, pembrolizumab) as new immunological molecules for MSI or MMRd (Table 1 and Figure 2).

**Platinum derivatives.** These compounds form covalent bonds with guanine and adenine in the DNA. The most important drugs in this group are cisplatin, carboplatin and oxaliplatin (Figure 2). Drugs containing platinum salts exert their cytotoxic effect by means of DNA adduct formation, leading to inhibition of DNA replication and apoptosis<sup>39</sup>. The major path of adduct elimination is the nucleotide excision repair (NER). During NER, damaged DNA and unwound DNA helices are identified by the action of several factors, including xeroderma pigmentosum proteins (XPD, XPC and XPA). Cleavages of the damaged DNA strand are performed by nucleases XPG (3') and ERCC1 (5'), and adducts are removed<sup>40</sup>.

The glutathione S-transferases (GSTs) are involved in the inactivation of platinum compounds, thus preventing cellular DNA damage and increasing the treatment efficacy<sup>41</sup>. Single nucleotide polymorphisms (SNPs) in GSTP1, GSTT1 and GSTM1 can alter GST

7

activity<sup>42</sup>. The decrease in enzyme activity has been linked to reduced detoxification capacity, leading to increased efficacy of platinum compounds. GSTP1 (Ile105Val) has been associated with reduced enzyme activity<sup>43</sup>, a deletion in GSTT1 leads to the absence of enzyme activity and a deletion in GSTM1 is linked with decreased survival rate<sup>44,45</sup>.

The excision repair cross-complementation (ERCC) is involved in nucleotide repair system<sup>46</sup>. Polymorphisms in genes encoding these repair proteins may contribute to interindividual differences to platinum toxicity. The association between toxicity and ERCC1 rs11615 has been studied in CRC<sup>47</sup>. The mutant T allele has been related to grade 1 neuropathy in oxaliplatin-treated patients, even though no association with a higher degree of neuropathy was observed. In addition, ERCC2 (XPD) is involved in the oxaliplatin pathway. rs13181 has been related with treatment effectiveness<sup>48</sup>. Meanwhile, ERCC4 rs1799801, ERCC5 rs2016073 and rs751402 are associated with platinum response<sup>49</sup>.

The X-ray repair cross-complementing protein (XRCC1) and its variant rs25487 are involved in the repair of broken DNA strands which can be induced by platinum compounds; such repair is carried out by an excision repair system<sup>50</sup>. It has been suggested that a deterioration in the efficiency of DNA repair caused by Gln's allele leads to greater efficacy of oxaliplatin<sup>50</sup>. On the other hand, the presence of XRCC3 rs1799794 has been associated with an increased risk of neutropenia<sup>51</sup>. Biomarkers focused on oxaliplatin are listed in Table  $2^{42,48-50,52-55}$ .

Pharmacogenomics identifies mutations that may predict an efficient therapeutic response; however, genetic variations significantly change among race/ethnic populations worldwide<sup>56</sup>. The allele frequencies of platinum derivative variants rs1695, rs11615, rs13181, rs1799801, rs2016073, rs2234671, rs25487 and rs1799794, according to the 1000 Genomes Project (phase 3), are shown in Table 3<sup>57</sup>.

Antimetabolites. These drugs inhibit enzymes related to purine and pyrimidine synthesis, resulting in cell depletion and alteration of nucleic acid synthesis. Among these, there are pyrimidine analogs such as 5-fluorouracil (5-FU) and oral pro-drugs such as gemcitabine, capecitabine and tegafur (Figure 2)<sup>58</sup>.

Fluoropyrimidines (5-FU, capecitabine and tegafur) are antimetabolite drugs used in CRC treatment. 5-FU is a fluoropyrimidine derivative with two major mechanisms of action that explain its cytotoxic effect<sup>59</sup>. The main active metabolite of 5-FU (5-FdUMP) prevents DNA synthesis by forming a complex with thymidylate synthase (TS) stabilized by 5,10-methylenetetrahydrofolate (5,10-MTHF), thus inhibiting the conversion of monophosphate 2'-deoxyuridine-5' (dUMP) to deoxythymidine-2'-5'-monophosphate (dTMP), an essential precursor for DNA synthesis. In addition, the incorporation of 5-FU to nucleotides in DNA and RNA strands leads to an alteration in the processing of nucleic acids<sup>59</sup>. Gemcitabine is a structural analog of deoxycytidine, which is metabolized by nucleoside kinase to nucleoside diphosphate and triphosphate<sup>60</sup>.

TYMS protein is a homodimeric methyltransferase enzyme that catalyzes the synthesis reaction of thymidylate. This reaction is a critical step in the formation of deoxythymidine 5'-triphosphate (dTTP), an indispensable metabolite in DNA synthesis. TYMS contains a

tandem of a polymorphic 28-base-pair sequence repeated in the promoter (TSER) 5' untranslated region  $(5'-UTR)^{42,60}$ .

Inactivation of 5-FU depends on dihydropyrimidine dehydrogenase (DPYD) activity<sup>61</sup>. The deficient activity of DPYD leads to prolonged 5-FU plasma half-life, causing a severe hematological toxicity<sup>62</sup>. DPYD deficiency is present in ~3% of all cancer patients, but it represents approximately 50% of patients manifesting severe toxicity. So far, ~30 polymorphisms in DYPD have been identified; however, a mutation of G>A at the splicing site in exon 14 (IVS14+1G>A) leads to the formation of a truncated protein without residual activity<sup>63</sup>. The incidence of this allele is rare, with a heterozygote population frequency of 0.9-1.8%. Nevertheless, it is estimated to be responsible for approximately 25% of all cases of 5-FU unexpected toxicity<sup>59</sup>.

Methylenetetrahydrofolate reductase (MTHFR) catalyzes the conversion of 5,10-MTHF to 5-MTHF. The most common polymorphisms of MTHFR are C677T and A1298C. These polymorphisms lead to decreased enzyme activity, which induces a more effective stabilization of the FdUMP-TS ternary complex, potentiating 5-FU toxicity<sup>64–66</sup>.

ATP-Binding Cassette Sub-Family B1 (ABCB1) is a member of the ABC transporter superfamily and its protein is known as P-glycoprotein<sup>67</sup>. ABCB1 overexpression in tumors has been associated with resistance to chemotherapeutic drugs. ABCB1 is a highly polymorphic gene that significantly differs among ethnic groups. Some of the most studied SNPs are rs1128503, rs2032592 and rs1045642<sup>68</sup>.

Cytidine deaminase (CDA) is involved in capecitabine metabolism in the liver to form 5fluorodeoxyuridine, which, in turn, becomes 5-FU by the action of thymidine phosphorylase (TP)<sup>60</sup>. The decreased activity of CDA leads to the accumulation of potentially toxic metabolites. The variation in expression of CDA has been linked to polymorphisms in the promoter region of CDA and affects the metabolism of gemcitabine and capecitabine<sup>67</sup>. rs602950 and rs532545 have been associated with increased expression of CDA *in vitro* in capecitabine-treated patients<sup>67</sup>.

TP is involved in 5-FU metabolism, where 5-FU is converted to 5-fluoro-2'-deoxyuridine  $(FUDR-5)^{69}$ . rs11479 generates an amino acid change of serine to leucine that leads to a lower treatment response. Enolase superfamily member 1 (ENOSF1) gene encodes an antisense RNA against TYMS. ENOSF1 regulates mRNA and protein expression of TYMS. Hence, TYMS variants with lower or higher activity affect its function<sup>63,64</sup>. Lastly, biomarkers focused on capecitabine, 5-FU and gemcitabine drugs are listed in Table  $4^{63,64,67-73}$ .

The allele frequencies of rs3918290, rs1801133, rs1128503, rs2072671, rs9344, rs9344 and rs2612091 polymorphisms in populations worldwide are shown in Table 5<sup>57</sup>.

**Agents interacting with topoisomerases.** Topoisomerases play a key role in the cell replication, transcription, and DNA repair. It modifies the tertiary DNA structure without altering the nucleotide sequence. In humans, three types of topoisomerases (I, II and III) have been identified. Within this group, camptothecin derivatives are included such as irinotecan<sup>74</sup> (Figure 2).

Irinotecan is a potent inhibitor of topoisomerase I<sup>75</sup>. It promotes an oxidation bioreaction mediated by CYP3A to form APC, a cytotoxic substance. Alternatively, irinotecan is converted by hepatic carboxylesterase to SN-38. This compound is conjugated further by several UDP glucuronyl to reach the inactive metabolite SN-38G<sup>76</sup>. To enable excretion, SN-38 and irinotecan are actively transported out of the cell by ATP-dependent efflux pump (ABCB1). After biliary excretion, SN-38G can become active SN-38 by bacterial beta-glucuronidase, which can lead to gastrointestinal toxicity.

It has been shown that reduced glucuronidation of SN-38 significantly increases irinotecan gastrointestinal toxicity<sup>77</sup>. The main UDP-glucuronosyltransferase (UGT) involved in conjugating SN-38 is UGT1A1. At least 25 UGT1A1 polymorphisms have been described, of which the most common in the promoter region consists of seven TA-repetitions (-53 [TA] 6>7, UGT1A1\*28) instead of six<sup>78,79</sup>. The highest number of TA repeats is associated with a reduction of UGT1A1 expression, leading to reduced glucuronidation. UGT1A1\*28 has proven to be a significant predictor of severe toxicity following administration of irinotecan<sup>80,81</sup>.

ABC transporters, including ABCC1, ABCC2, ABCB1 and ABCG2, regulate output of hepatic and biliary CPT-11 metabolites<sup>55,82</sup>. SNPs in ABCB1 and ABCC2 have been recently associated with modulation of CPT-11 and SN-38 exposure<sup>75</sup>. In addition, other SNPs in ABCC5 and ABCG2 genes have been correlated with both hematological and non-hematological toxicities<sup>83</sup>.

The solute carrier organic anion transporter family member 1B1 (SLCO1B1) is an important transporter expressed in the basolateral membrane of hepatocytes which mediates the availability of active irinotecan metabolite<sup>84</sup>. rs4149056 has been associated with an increased SN-38 concentration in patients with metastatic CRC (mCRC). Meanwhile, other polymorphisms are linked to faster response rate and higher PFS<sup>76</sup>. Finally, biomarkers focused on irinotecan drug are shown in Table  $6^{45,52,68,78,79,81,83-89}$ .

The allele frequencies of genes that interact with topoisomerases rs2244613, rs1045642, rs2074087, rs262604, rs1051266 and rs2306283 in populations worldwide are shown in Table  $7^{57}$ .

Antiangiogenics. Vascular endothelial growth factor (VEGF) is a major regulator of angiogenesis and inhibition mediated by bevacizumab which reduces tumor volume<sup>90,91</sup>. Bevacizumab is a recombinant humanized monoclonal IgG antibody directed against all isoforms of VEGF-A. Hypoxia is a potent stimulus for VEGF expression and one of the control elements in this mechanism is the hypoxia-inducible factor 1A (HIF-1a)<sup>92</sup>. This factor binds to a 28-bp promoter in the 5' upstream region of VEGF, thereby/hence-stimulating transcription. In addition, other regulatory elements for VEGF expression are found in the 3'-UTR as shown in Table 8<sup>92–95</sup>.

Variations in the VEGF receptor 1 (rs9582036) and 2 (rs12505758) are associated with tyrosine kinase domain. High expression levels of these receptors contribute to a less favorable outcome when treated with bevacizumab, associated with PFS and overall survival  $(OS)^{90,94,96}$ .

Various studies from mCRC have investigated the predictive impact of some SNPs present in VEGF-A, which are involved in bevacizumab response. Loupakis *et al.*, conducted a retrospective analysis which found that rs833061 was associated with PFS and OS<sup>94</sup>. Meanwhile, Sibertin-Blanc *et al.*, showed that T-carriers of the C237T SNP had shorter time-to-treatment failure as well as shorter PFS and OS<sup>97</sup>.

Annexin A11 (ANXA11) has been associated with a spectrum of regulatory functions in calcium signaling, cell division and apoptosis<sup>98</sup>. ANXA11 rs1049550 leads to an amino acid change (R230C) of the first conserved domain of annexin, which is responsible for  $Ca^{+2}$  dependent intracellular traffic. Response to bevacizumab revealed that patients carrying rs1049550 were more sensitive to chemotherapy than those having at least one C allele<sup>95</sup>.

CXC chemokine receptors (CXCR1 and CXCR2) are integral membrane proteins which specifically bind and respond to CXC chemokine family cytokines<sup>93</sup>. They represent a family of seven receptors linked to G-protein that plays an important role in angiogenesis. CXCR1 rs2234671 and CXCR2 rs2230054 are associated with overall response (ORR)<sup>99</sup>.

Finally, the Food and Drug Administration (FDA) and the European Medicines Agency (EMA) have approved new molecules that improve the therapeutic effectiveness, OS and PFS. In particular, research efforts have focused on novel agents targeting tumor angiogenic activity, cell growth and migration in mCRC. The use of molecules targeting

VEGF pathways (ziv-aflibercept, regorafenib and ramucirumab) have been integrated into clinical practice<sup>90</sup> (Figure 2).

The allele frequencies of variants that interact with antiangiogenic agents rs9582036, rs12505758, rs3025039, rs1049550 and rs2230054 in populations worldwide are shown in Table  $9^{57}$ .

**Agents against epidermal growth factor receptor.** Cetuximab and panitumumab are monoclonal antibodies (mAb) that block the action of EGF and may be employed in mCRC treatment<sup>7,92</sup>. These drugs exert their action by binding to the extracellular domain of EGFR, with a greater affinity than the wild-type EGF, thereby blocking phosphorylation induced by EGFR ligands. Some variants are shown in Table 10<sup>62,63,73,100–108</sup>.

The EGF/EGFR pathway plays an important role in cancer pathogenesis. EGF and EGFR are commonly overexpressed in CRC and they appear to be associated with poor prognosis and increased metastatic risk<sup>109</sup>. EGFR is a transmembrane glycoprotein that plays a main role in cell proliferation, migration and survival. EGFR R497K attenuates tyrosine kinase activation<sup>110</sup>, and EGF G61A increases its production when individuals have GG or GA genotypes<sup>105</sup>. The EGF/EGFR pathway is a predictive marker for cetuximab treatment in patients with locally advanced CRC<sup>111</sup>.

KRAS oncogene is a member of the human RAS family, which produces a self-inactivating guanosine triphosphate (GTP), binding signal transducer located on the inner surface of the cell membrane<sup>112</sup>. KRAS mutations may compromise the intrinsic GTPase activity,

resulting in constitutively active KRAS protein that affects various signaling pathways<sup>112</sup>. The 45% of CRC cases has KRAS mutations and it has been shown that these mutations are predictive biomarkers of poor outcome in mCRC treated with cetuximab<sup>77</sup>. The anti-EGFR mAb therapy significantly improves both PFS and OS tumors without mutations in RAS. Therefore, mutations in KRAS predict resistance to mAb directed to EGFR with cetuximab and panitumumab<sup>100</sup>. NRAS codifies an isoform of RAS protein, involved primarily in the regulation of cell division<sup>108</sup>. Mutations in exon 2, 3 and 4 of NRAS, in addition to those in exon 2 of KRAS, must be detected before administration of a monoclonal anti-EGFR<sup>100</sup>. According to the NCCN, KRAS and NRAS are the only one predictive biomarkers approved in mCRC. Cetuximab and panitumumab are applied on patients with non-mutated RAS, and bevacizumab is applied on patients with mutated RAS<sup>12,113</sup>.

Regarding Fc receptor range, modulating the immune response could be a further important mechanism to cetuximab sensitivity. The immune mechanism of antibody-dependent (ADCC) mediated cellular cytotoxicity through Fc receptors (Fc gamma R) made by immune cells, plays an important role in the effect of IgG1 antitumor antibodies<sup>114,115</sup>. The most common polymorphisms in FCGR2A and FCGR3A are rs1801274 and rs396991, respectively<sup>116,117</sup>.

The phosphatidylinositol 3-kinase (PI3K/AKT) pathway plays an essential role in cancer pathogenesis, being imperative the design of therapeutic inhibitors<sup>118</sup>. PIK3CA gene encodes the p110 catalytic subunit of PI3K alpha. PIK3CA mutations (associated with KRAS mutation and MSI) stimulate AKT pathway and promote cell growth in CRC<sup>118</sup>.

PIK3CA mutations in exon 9 and 20 affects the helical and kinase domains of the protein, promoting a lack of effectiveness in drug treatments<sup>106</sup>.

BRAF protein is part of the RAS/MAPK signaling pathway, which regulates cell growth, proliferation, migration and apoptosis<sup>119</sup>. BRAF is a driver gene whose mutations are inversely associated with treatment response and are mutually exclusive with RAS mutations<sup>108,120</sup>. Hence, BRAF V600E mutation correlates with worse prognosis<sup>121</sup>. Vemurafenib is a third-line therapeutic option in advanced mCRC with BRAF mutations<sup>122</sup>. Furthermore, it has been proposed that patients with KRAS and BRAF mutations could be eligible for mAb treatment against EGFR. Finally, BRAF must not present any mutation for a favorable treatment response when panitumumab or cetuximab are applied<sup>107</sup>.

COX is the limiting enzyme in the conversion of arachidonic acid into prostaglandins. COX2, encoded by prostaglandin endoperoxide synthase 2 (PTGS2), is involved in metastasis and chemotherapy resistance<sup>123</sup>. High levels of COX2 are associated with shorter OS in CRC. The C allele of COX2 G765C polymorphism has been associated with a significantly lower promoter activity<sup>104,124</sup>.

The allele frequencies of rs2227983, rs4444903, rs1801274, rs396991 and rs20417 genetic variants in populations worldwide are shown in Table  $11^{57}$ .

**Colorectal cancer immunogenomics.** Recent advances in cancer immunology have highlighted the immunogenic nature of CRC and provided insights regarding the complex tumor-immune system interactions that drive immune evasion in  $CRC^{125,126}$ . One of the

mechanisms that mediates tumor-associated immune escape is the activation of inhibitory co-receptors or immune checkpoints on the T lymphocyte surface by tumor cells through the expression of immunosuppressive molecules<sup>125–127</sup>.

Programmed cell death protein 1 (PD-1, CD279) is an inhibitory co-receptor expressed by exhausted tumor-infiltrating lymphocytes (TILs) present within the tumor microenvironment<sup>127–129</sup>. PD-1 engages with programmed-death ligands 1 (PD-L1, BT-H1, CD274) and 2 (PD-L2, B7-DC, CD273) which are expressed by CRC cells<sup>130–133</sup>. PD-1/PD-L1 interaction inhibits CD8<sup>+</sup>T-cell activation, cytokine production, proliferation and cytotoxicity which suppresses the host immune response and allows CRC cells to proliferate and metastasize<sup>127–129</sup>.

Immune checkpoint inhibition has revolutionized cancer immunotherapy since it has proven to be very successful for treatment of melanoma and non-small cell lung cancer<sup>134–136</sup>. It has been shown that PD-1 blockade is a highly efficient therapeutic strategy against MSI-high and MMRd CRC tumors since these tumors display dense lymphocyte infiltrates due to their increased expression of immunogenic neo-antigens<sup>137–139</sup>. Moreover, these tumors exhibit higher PD-1 expression on TILs and PD-L1 expression than microsatellite stable tumors<sup>139,140</sup>. Subsequently, the FDA approved pembrolizumab and nivolumab, two anti PD-1 antibodies, for treatment of metastatic MSI-high or MMRd solid tumors.

According to Tauriello *et al.*, inhibition of the PD-1 and PD-L1 immune checkpoints provoked a limited response in quadruple-mutant mice. By contrast, his results strongly suggest that inhibition of TGF $\beta$  signaling could be promising as immunotherapy for patients with microsatellite stability and stroma-rich CRCs, enduring cytotoxic T-cell response against tumor cells that prevent metastasis<sup>141–143</sup>. The clinical implications of CRC immunogenomics continue to expand, and it will likely serve as a guide for next-generation immunotherapy strategies for improving outcomes for this disease (Figure 2).

### The Pan-Cancer Atlas: germline pathogenic variants

The Pan-Cancer Atlas provides a panoramic view of the oncogenic processes that contribute to human cancer. It reveals how genetic variants collaborate in cancer progression and explores the influence of mutations on cell signaling and immune cell composition, providing insight to prioritize the development of new immunotherapies<sup>144</sup>.

According to Huang *et al.*, the Pan-Cancer Atlas analyzed 564 CRC samples and found several pathogenic germline variants in the APC, ATM, ATR, BARD1, BLM, BRCA1, BRCA2, BRIP1, CHEK2, COL7A1, FANCI, GJB2, MLH1, MSH2, MSH6, PALB2, POT1, RAD51D, RECQL4, RET, RHBDF2 and SDHA genes<sup>145</sup>. Additionally, Table 12 shows the allele frequencies of those 29 pathogenic germline variants according to The Exome Aggregation Consortium (ExAC)<sup>146</sup>.

### Biomarker network in colorectal cancer

Figure 3 shows the proposed biomarker network in CRC. The protein-protein interaction (PPi) network with a highest confidence cutoff of 0.9 was created using String Database<sup>147</sup>.

This network is made up of known and predicted interactions of driver genes, genes with pathogenic germline mutations according to the Pan-Cancer Atlas<sup>33,145</sup>, genes with somatic mutations according to the CGI<sup>34</sup>, and druggable enzymes according to the Pharmacogenomics Knowledge Base (PharmGKB)<sup>148,149</sup>.

The enrichment analysis of gene ontology (GO) terms related to biological processes and metabolic pathways were carried in the 87 genes of CRC biomarker network (Figure 3). The top biological processes with significant false discovery rate (FDR) <0.01 were DNA synthesis involved in DNA repair, strand displacement and response to drug. Meanwhile, the top metabolic pathways with FDR <0.01 were colorectal, endometrial and pancreatic cancer types<sup>150</sup>.

### Pharmacogenomics in clinical practice

In addition to the NCCN and ESMO guidelines<sup>36–38</sup>, the Canadian Pharmacogenomics Network for Drug Safety (CPNDS), the Royal Dutch Association for the Advancement of Pharmacy (DPWG) and the Clinical Pharmacogenetics Implementation Consortium (CPIC) have published precise guidelines for the application of pharmacogenomics in clinical practice<sup>151–153</sup>. All this information is published in the PharmGKB which is a comprehensive resource that curates knowledge about the impact of 80 clinical annotations on drug response<sup>148</sup> (Supplementary Table 2).

In addition to the 1000 Genomes Project (Phase 3)<sup>57</sup>, there are 33 studies that have published the allele frequencies of mutations in several druggable enzymes in 34 different

ethnic populations from Latin America (Supplementary Tables 3-10). Figure 4 is an innovative way to visualize and correlate the minor allele frequencies of 43 genes related to the different categories of drugs applied in CRC treatments in 8674 samples from 9 Latin American countries. This information will make it easier for clinical oncologists to make decisions regarding CRC treatments. For instance, the minor allele (G) of GSTP1 rs1695 can alter GST activity reducing detoxification capacity and leading to increased efficacy of platinum compounds<sup>42,43</sup>. Thus, the Latin American countries whom best reduce the detoxification capacity and increase the efficacy of platinum compounds are Venezuela, Mexico and Peru due to their populations have a G allele frequency  $\geq 0.50$ .

On the other hand, it is imperative to unify efforts at the governmental level to increase investment in pharmacogenomics fomenting precision medicine in clinical practice. The most relevant barriers to implement pharmacogenomics testing in clinical practice in Latin America are: 1) need for clear guidelines for the use of pharmacogenomics, 2) insufficient awareness of pharmacogenomics among clinicians, and 3) absence of a regulatory institution that facilitates the use of pharmacogenomics tests<sup>154</sup>. Overcoming the previously mentioned barriers, pharmacogenomics will make it possible to improve public health investment, patient safety and drug dosage in CRC treatments<sup>155</sup>.

#### CONCLUSION

In the era of precision medicine, it is important to unify all current knowledge about the CRC biology to improve patient treatments. Large-scale projects worldwide have studied the multi-omics landscape of CRC by implementing the CMS classification and generating

new therapeutic targets related to different populations worldwide. Developed countries might incorporate racial/ethnic minority populations in future cancer researches and clinical trials, and developing countries might invest to obtain a database of genomic profiles of their populations with the overall objective of linking pharmacogenomics in clinical practice.

## Acknowledgements

This research was supported by Centro de Investigación Genética y Genómica of Universidad UTE and the Latin American Society of Pharmacogenomics and Personalized Medicine.

## Author contributions

ALC and NWS conceived the subject and wrote the manuscript. ALC, GJK, DPIB, JGC, IAC and PGR did data curation and supplementary data. SG, CPyM, PEL, LAQ and JPC made a substantial contribution to the discussion of content. All authors reviewed and/or edited the article before submission.

### **Competing interests**

The authors declare no competing interest.

## Data availability statement

All data generated or analysed during this study are included in this published article (and its Supplementary Information).

## Disclaimer

The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health.

# REFERENCES

- 1. Dienstmann, R. *et al.* Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat. Rev. Cancer* **17**, 79–92 (2017).
- Puccini, A. & Lenz, H.-J. Colorectal cancer in 2017: Practice-changing updates in the adjuvant and metastatic setting. *Nat. Rev. Clin. Oncol.* 2–3 (2017). doi:10.1038/nrclinonc.2017.185
- 3. Sabarinathan, R. *et al.* The whole-genome panorama of cancer drivers. *bioRxiv* 190330 (2017). doi:10.1101/190330
- 4. López-Cortés, A., Guerrero, S., Redal, M. A., Alvarado, A. T. & Quiñones, L. A. State of art of cancer pharmacogenomics in Latin American populations. *International Journal of Molecular Sciences* **18**, (2017).
- 5. Gupta, P. D. Pharmacogenetics, pharmacogenomics and ayurgenomics for personalized medicine: a paradigm shift. *Indian J. Pharm. Sci.* **77**, 135–41 (2015).
- 6. Wilke, R. A. *et al.* Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. *Nat. Rev. Drug Discov.* **6**, 904–16 (2007).
- 7. Sadanandam, A. *et al.* A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat. Med.* **19**, 619–625 (2013).
- 8. De Sousa E Melo, F. *et al.* Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat. Med.* **19,** 614–618 (2013).
- 9. Perez-Villamil, B. *et al.* Colon cancer molecular subtypes identified by expression profiling and associated to stroma, mucinous type and different clinical behavior. *BMC Cancer* **12**, 260 (2012).
- 10. Vogelstein, B. *et al.* Genetic Alterations during Colorectal-Tumor Development. *N. Engl. J. Med.* **319**, 525–532 (1988).
- Marisa, L. *et al.* Gene Expression Classification of CRC into Molecular Subtypes: Characterization, Validation, and Prognostic Value. *PLoS Med.* **10**, e1001453 (2013).
- 12. Strimpakos, A. S., Syrigos, K. N. & Saif, M. W. Pharmacogenetics and biomarkers in colorectal cancer. *Pharmacogenomics J.* **9**, 147–160 (2009).
- 13. Van Cutsem, E. *et al.* Randomized phase III trial comparing biweekly infusional fluorouracil/leucovorin alone or with irinotecan in the adjuvant treatment of stage III colon cancer: PETACC-3. *J. Clin. Oncol.* **27**, 3117–25 (2009).
- 14. Van Dongen, S. Graph Clustering Via a Discrete Uncoupling Process. *SIAM J. Matrix Anal. Appl.* **30**, 121–141 (2008).

- 15. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for largescale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).
- Marisa, L. *et al.* Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value. *PLoS Med.* 10, (2013).
- 17. Hoadley, K. A. *et al.* Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **158**, 929–944 (2014).
- Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* 21, 1350–1356 (2015).
- 19. Guerrero, S. *et al.* Analysis of Racial/Ethnic Representation in Select Basic and Applied Cancer Research Studies. *Sci. Rep.* **8**, 13978 (2018).
- 20. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–7 (2012).
- Roepman, P. *et al.* Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. *Int. J. Cancer* 134, 552–562 (2013).
- 22. Budinska, E. *et al.* Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J. Pathol.* **231**, 63–76 (2013).
- 23. Schlicker, A. *et al.* Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. *BMC Med. Genomics* **5**, 66 (2012).
- Isella, C. *et al.* Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nat. Commun.* 8, 1–16 (2017).
- 25. Rodriguez-Salas, N. *et al.* Clinical relevance of colorectal cancer molecular subtypes. *Critical Reviews in Oncology/Hematology* (2017). doi:10.1016/j.critrevonc.2016.11.007
- 26. Calon, A. *et al.* Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47**, 320–329 (2015).
- Widder, M. *et al.* Multipotent mesenchymal stromal cells promote tumor growth in distinct colorectal cancer cells by a β1-integrin-dependent mechanism. *Int. J. Cancer* (2015). doi:10.1002/ijc.29844
- 28. Tamborero, D. *et al.* Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci. Rep.* **3**, 2650 (2013).
- 29. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. *Nature* **300**, 149–152 (1982).

- 30. Tabin, C. J. *et al.* Mechanism of activation of a human oncogene. *Nature* **300**, 143–9 (1982).
- 31. Futreal, P. A. *et al.* A census of human cancer genes. *Nat. Rev. Cancer* **4**, 177–183 (2004).
- 32. Rubio-Perez, C. *et al.* In Silico Prescription of Anticancer Drugs to Cohorts of 28 Tumor Types Reveals Targeting Opportunities. *Cancer Cell* **27**, 382–396 (2015).
- 33. Bailey, M. H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173**, 371–385.e18 (2018).
- 34. Tamborero, D. *et al.* Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. 1–8 (2018). doi:10.1186/s13073-018-0531-8
- 35. Sondka, Z. *et al.* The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* 1 (2018). doi:10.1038/s41568-018-0060-1
- 36. Van Cutsem, E. *et al.* ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann. Oncol.* **0**, mdw235 (2016).
- Yoshino, T. *et al.* Pan-Asian adapted ESMO consensus guidelines for the management of patients with metastatic colorectal cancer; A JSMO - ESMO initiative endorsed by CSCO, KACO, MOS, SSO and TOS. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 2–70 (2017). doi:10.1093/annonc/mdx738
- National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Colon Cancer. *Natl. Compr. Cancer Netw.* 1–180 (2018). doi:10.1093/med/9780199664535.003.0012
- 39. Mehmood, R. K. Review of Cisplatin and oxaliplatin in current immunogenic and monoclonal antibody treatments. *Oncol. Rev.* **8**, 256 (2014).
- 40. Friedberg, E. C. Nucleotide Excision Repair in Eukaryotes. *eLS* **36**, 1–8 (2001).
- 41. McWhinney, S. R., Goldberg, R. M. & McLeod, H. L. Platinum neurotoxicity pharmacogenetics. *Mol. Cancer Ther.* **8**, 10–6 (2009).
- 42. Kumamoto, K. *et al.* Polymorphisms of GSTP1, ERCC2 and TS-3'UTR are associated with the clinical outcome of mFOLFOX6 in colorectal cancer patients. *Oncol. Lett.* **6**, 648–654 (2013).
- 43. Inada, M. *et al.* Associations between oxaliplatin-induced peripheral neuropathy and polymorphisms of the ERCC1 and GSTP1 genes. *Int. J. Clin. Pharmacol. Ther.* **48**, 729–34 (2010).
- 44. Kap, E. J. *et al.* Genetic variants in the glutathione S-transferase genes and survival in colorectal cancer patients after chemotherapy and differences according to treatment with oxaliplatin. *Pharmacogenet. Genomics* **24**, 340–7 (2014).

- 45. McLeod, H. L. *et al.* Pharmacogenetic predictors of adverse events and response to chemotherapy in metastatic colorectal cancer: results from North American Gastrointestinal Intergroup Trial N9741. *J. Clin. Oncol.* **28**, 3227–33 (2010).
- 46. Takenaka, T. *et al.* Effects of excision repair cross-complementation group 1 (ERCC1) single nucleotide polymorphisms on the prognosis of non-small cell lung cancer patients. *Lung Cancer* **67**, 101–7 (2010).
- 47. Ma, S.-C., Zhao, Y., Zhang, T., Ling, X.-L. & Zhao, D. Association between the ERCC1 rs11615 polymorphism and clinical outcomes of oxaliplatin-based chemotherapies in gastrointestinal cancer: a meta-analysis. *Onco. Targets. Ther.* **8**, 641–8 (2015).
- 48. Yin, M. *et al.* ERCC1 and ERCC2 polymorphisms predict clinical outcomes of oxaliplatin-based chemotherapies in gastric and colorectal cancer: a systemic review and meta-analysis. *Clin. Cancer Res.* **17**, 1632–40 (2011).
- 49. Chen, J. *et al.* ERCC5 promoter polymorphisms at -763 and +25 predict the response to oxaliplatin-based chemotherapy in patients with advanced colorectal cancer. *Cancer Biol. Ther.* **8**, 1424–30 (2009).
- 50. Zaanan, A. *et al.* ERCC1, XRCC1 and GSTP1 Single Nucleotide Polymorphisms and Survival of Patients with Colon Cancer Receiving Oxaliplatin-Based Adjuvant Chemotherapy. *J. Cancer* **5**, 425–32 (2014).
- 51. Di Francia, R. *et al.* Current strategies to minimize toxicity of oxaliplatin: selection of pharmacogenomic panel tests. *Anti-cancer drugs* **24**, 1069–78 (2013).
- 52. Ruzzo, A. *et al.* Genetic markers for toxicity of adjuvant oxaliplatin and fluoropyrimidines in the phase III TOSCA trial in high-risk colon cancer patients. *Sci. Rep.* **4**, (2014).
- 53. Sullivan, I. *et al.* Pharmacogenetics of the DNA repair pathways in advanced nonsmall cell lung cancer patients treated with platinum-based chemotherapy. *Cancer Lett.* **353**, 160–6 (2014).
- 54. Artac, M. *et al.* The value of XPD and XRCC1 genotype polymorphisms to predict clinical outcome in metastatic colorectal carcinoma patients with irinotecan-based regimens. *J. Cancer Res. Clin. Oncol.* **136**, 803–9 (2010).
- Cecchin, E. *et al.* A prospective validation pharmacogenomic study in the adjuvant setting of colorectal cancer patients treated with the 5-fluorouracil/leucovorin/oxaliplatin (FOLFOX4) regimen. *Pharmacogenomics J.* 13, 403–9 (2013).
- 56. Huang, T., Shu, Y. & Cai, Y.-D. Genetic differences among ethnic groups. *BMC Genomics* **16**, 1093 (2015).
- 57. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

- 58. Walko, C. M. & Lindley, C. Capecitabine: a review. *Clin. Ther.* 27, 23–44 (2005).
- 59. Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330–8 (2003).
- 60. Loganayagam, A. *et al.* Pharmacogenetic variants in the DPYD, TYMS, CDA and MTHFR genes are clinically significant predictors of fluoropyrimidine toxicity. *Br. J. Cancer* **108**, 2505–15 (2013).
- 61. Toffoli, G. *et al.* Clinical validity of a DPYD-based pharmacogenetic test to predict severe toxicity to fluoropyrimidines. *Int. J. cancer* **137**, 2971–80 (2015).
- 62. Galván, C. A., Elbarcha, O. C., Fernández, E. J., Beltramo, D. M. & Soria, N. W. Genetic profiling of GSTP1, DPYD, FCGR2A, FCGR3A and CCND1 genes in an Argentinian population. *Clin. Biochem.* **44**, 1058–1061 (2011).
- 63. Koumarianou, A. *et al.* Prognostic markers in early-stage colorectal cancer: significance of TYMS mRNA expression. *Anticancer Res.* **34**, 4949–62 (2014).
- 64. Rosmarin, D. *et al.* A candidate gene study of capecitabine-related toxicity in colorectal cancer identifies new toxicity variants at DPYD and a putative role for ENOSF1 rather than TYMS. *Gut* **64**, 111–20 (2015).
- 65. López-Cortés, A. *et al.* Breast cancer risk associated with gene expression and genotype polymorphisms of the folate-metabolizing MTHFR gene: a case-control study in a high altitude Ecuadorian mestizo population. *Tumour Biol.* **36**, 6451–61 (2015).
- 66. Funke, S., Brenner, H. & Chang-Claude, J. Pharmacogenetics in colorectal cancer: A systematic review. *Pharmacogenomics* **9**, 1079–1099 (2008).
- 67. García-González, X. *et al.* Variants in CDA and ABCB1 are predictors of capecitabine-related adverse reactions in colorectal cancer. *Oncotarget* **6**, 6422–30 (2015).
- 68. Gonzalez-Haba, E. *et al.* ABCB1 gene polymorphisms are associated with adverse reactions in fluoropyrimidine-treated colorectal cancer patients. *Pharmacogenomics* **11**, 1715–23 (2010).
- 69. Huang, L. *et al.* Thymidine phosphorylase gene variant, platelet counts and survival in gastrointestinal cancer patients treated by fluoropyrimidines. *Sci. Rep.* **4**, 5697 (2014).
- 70. Jennings, B. A. *et al.* Functional polymorphisms of folate metabolism and response to chemotherapy for colorectal cancer, a systematic review and meta-analysis. *Pharmacogenet. Genomics* **22**, 290–304 (2012).
- 71. Teh, L. K. *et al.* Potential of dihydropyrimidine dehydrogenase genotypes in personalizing 5-fluorouracil therapy among colorectal cancer patients. *Ther. Drug Monit.* **35**, 624–30 (2013).

- 72. Afzal, S. *et al.* MTHFR polymorphisms and 5-FU-based adjuvant chemotherapy in colorectal cancer. *Ann. Oncol.* **20**, 1660–1666 (2009).
- 73. Absenger, G. *et al.* The cyclin D1 (CCND1) rs9344 G> A polymorphism predicts clinical outcome in colon cancer patients treated with adjuvant 5-FU-based chemotherapy. *Pharmacogenomics J.* **14**, 130–134 (2014).
- 74. Armand, J. P. *et al.* CPT-11 (irinotecan) in the treatment of colorectal cancer. *Eur. J. Cancer* **31A**, 1283–7 (1995).
- 75. Ando, Y. & Hasegawa, Y. Clinical pharmacogenetics of irinotecan (CPT-11). *Drug Metab. Rev.* **37**, 565–74 (2005).
- 76. Innocenti, F. *et al.* Comprehensive pharmacogenetic analysis of irinotecan neutropenia and pharmacokinetics. *J. Clin. Oncol.* **27**, 2604–14 (2009).
- 77. Fiala, O. *et al.* High serum level of C-reactive protein is associated with worse outcome of patients with advanced-stage NSCLC treated with erlotinib. *Tumour Biol.* **36**, 9215–22 (2015).
- Liu, X. & Xu, W. UGT1A1\*28 polymorphisms: a potential pharmacological biomarker of irinotecan-based chemotherapies in colorectal cancer. *Pharmacogenomics* 15, 1171–4 (2014).
- 79. Maeda, H. *et al.* Differences in UGT1A1, UGT1A7, and UGT1A9 polymorphisms between Uzbek and Japanese populations. *Mol. Diagn. Ther.* **18**, 333–42 (2014).
- 80. Yao, L. *et al.* The association between TA-repeat polymorphism in the promoter region of UGT1A1 and breast cancer risk: a meta-analysis. *Breast Cancer Res. Treat.* **122**, 879–882 (2010).
- 81. Cecchin, E. *et al.* Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan. *J. Clin. Oncol.* **27**, 2457–65 (2009).
- 82. Sai, K. *et al.* Additive effects of drug transporter genetic polymorphisms on irinotecan pharmacokinetics/pharmacodynamics in Japanese cancer patients. *Cancer Chemother. Pharmacol.* **66**, 95–105 (2010).
- 83. Cha, P.-C. *et al.* Single nucleotide polymorphism in ABCG2 is associated with irinotecan-induced severe myelosuppression. *J. Hum. Genet.* **54**, 572–80 (2009).
- 84. Huang, L. *et al.* SLCO1B1 and SLC19A1 gene variants and irinotecan-induced rapid response and survival: a prospective multicenter pharmacogenetics study of metastatic colorectal cancer. *PLoS One* **8**, e77223 (2013).
- 85. Kubo, T. *et al.* Functional characterization of three naturally occurring single nucleotide polymorphisms in the CES2 gene encoding carboxylesterase 2 (HCE-2). *Drug Metab. Dispos.* **33**, 1482–7 (2005).

- 86. Swami, U., Goel, S. & Mani, S. Therapeutic targeting of CPT-11 induced diarrhea: a case for prophylaxis. *Curr. Drug Targets* **14**, 777–97 (2013).
- 87. Marsh, S. & Hoskins, J. M. Irinotecan pharmacogenomics. *Pharmacogenomics* **11**, 1003–10 (2010).
- 88. Yamanaka, H. *et al.* A novel polymorphism in the promoter region of human UGT1A9 gene (UGT1A9\*22) and its effects on the transcriptional activity. *Pharmacogenetics* **14**, 329–32 (2004).
- 89. Teft, W. A. *et al.* OATP1B1 and tumour OATP1B3 modulate exposure, toxicity, and survival after irinotecan-based chemotherapy. *Br. J. Cancer* **112**, 857–65 (2015).
- Bignucolo, A., De Mattia, E., Cecchin, E., Roncato, R. & Toffoli, G. Pharmacogenomics of targeted agents for personalization of colorectal cancer treatment. *Int. J. Mol. Sci.* 18, (2017).
- 91. Zondor, S. D. & Medina, P. J. Bevacizumab: an angiogenesis inhibitor with efficacy in colorectal and other malignancies. *Ann. Pharmacother.* **38**, 1258–64 (2004).
- 92. Zhang, W. *et al.* Molecular predictors of combination targeted therapies (cetuximab, bevacizumab) in irinotecan-refractory colorectal cancer (BOND-2 study). *Anticancer Res.* **30**, 4209–17 (2010).
- 93. Brat, D. J., Bellail, A. C. & Van Meir, E. G. The role of interleukin-8 and its receptors in gliomagenesis and tumoral angiogenesis. *Neuro. Oncol.* **7**, 122–33 (2005).
- 94. Loupakis, F. *et al.* Prospective validation of candidate SNPs of VEGF/VEGFR pathway in metastatic colorectal cancer patients treated with first-line FOLFIRI plus bevacizumab. *PLoS One* **8**, e66774 (2013).
- 95. Kim, J. C. *et al.* Feasibility of proposed single-nucleotide polymorphisms as predictive markers for targeted regimens in metastatic colorectal cancer. *Br. J. Cancer* **108**, 1862–9 (2013).
- 96. Formica, V. *et al.* Predictive value of VEGF gene polymorphisms for metastatic colorectal cancer patients receiving first-line treatment including fluorouracil, irinotecan, and bevacizumab. *Int. J. Colorectal Dis.* **26**, 143–51 (2011).
- 97. Sibertin-Blanc, C. *et al.* Vascular Endothelial Growth Factor A c.\*237C>T polymorphism is associated with bevacizumab efficacy and related hypertension in metastatic colorectal cancer. *Dig. Liver Dis.* **47**, 331–337 (2015).
- 98. Wang, J. et al. Annexin A11 in disease. Clin. Chim. Acta. 431, 164–8 (2014).
- 99. Gerger, A. *et al.* Pharmacogenetic angiogenesis profiling for first-line Bevacizumab plus oxaliplatin-based chemotherapy in patients with metastatic colorectal cancer. *Clin. Cancer Res.* **17**, 5783–92 (2011).

- 100. Sorich, M. J. *et al.* Extended RAS mutations and anti-EGFR monoclonal antibody survival benefit in metastatic colorectal cancer: a meta-analysis of randomized, controlled trials. *Ann. Oncol.* **26**, 13–21 (2015).
- 101. Waring, P., Tie, J., Maru, D. & Karapetis, C. S. RAS Mutations as Predictive Biomarkers in Clinical Management of Metastatic Colorectal Cancer. *Clinical Colorectal Cancer* (2015). doi:10.1016/j.clcc.2015.10.006
- 102. Pander, J. *et al.* Correlation of FCGR3A and EGFR germline polymorphisms with the efficacy of cetuximab in KRAS wild-type metastatic colorectal cancer. *Eur. J. Cancer* **46**, 1829–34 (2010).
- Bonin, S. *et al.* A synonymous EGFR polymorphism predicting responsiveness to anti-EGFR therapy in metastatic colorectal cancer patients. *Tumour Biol.* (2015). doi:10.1007/s13277-015-4543-3
- 104. Lurje, G. *et al.* Polymorphisms in cyclooxygenase-2 and epidermal growth factor receptor are associated with progression-free survival independent of K-ras in metastatic colorectal cancer patients treated with single-agent cetuximab. *Clin. Cancer Res.* 14, 7884–95 (2008).
- 105. Lu, X. *et al.* Polymorphism in epidermal growth factor is related to clinical outcomes of metastatic colorectal cancer patients treated with cetuximab: a systematic review and meta-analysis. *Int. J. Clin. Exp. Med.* **8**, 10929–37 (2015).
- 106. Ogino, S. *et al.* PIK3CA mutation is associated with poor prognosis among patients with curatively resected colon cancer. *J. Clin. Oncol.* **27**, 1477–84 (2009).
- 107. Di Nicolantonio, F. *et al.* Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J. Clin. Oncol.* **26**, 5705–12 (2008).
- 108. Oikonomou, E., Koustas, E., Goulielmaki, M. & Pintzas, A. BRAF vs RAS oncogenes: are mutations of the same pathway equal? Differential signalling and therapeutic implications. *Oncotarget* **5**, 11752–77 (2014).
- 109. Normanno, N. *et al.* Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* **366**, 2–16 (2006).
- 110. Wieduwilt, M. J. & Moasser, M. M. The epidermal growth factor receptor family: biology driving targeted therapeutics. *Cell. Mol. Life Sci.* **65**, 1566–84 (2008).
- 111. Hu-Lieskovan, S. *et al.* EGF61 polymorphism predicts complete pathologic response to cetuximab-based chemoradiation independent of KRAS status in locally advanced rectal cancer patients. *Clin. Cancer Res.* **17**, 5161–9 (2011).
- 112. Baines, A. T., Xu, D. & Der, C. J. Inhibition of Ras for cancer treatment: the search continues. *Future Med. Chem.* **3**, 1787–808 (2011).
- 113. Lurie, R. H., Cancer, C. & Cooper, H. S. Colon Cancer. (2018). doi:10.1093/med/9780199664535.003.0012

- 114. Alderson, K. L. & Sondel, P. M. Clinical cancer therapy by NK cells via antibodydependent cell-mediated cytotoxicity. *J. Biomed. Biotechnol.* **2011**, 379123 (2011).
- 115. Calemma, R. *et al.* Fc gamma receptor IIIa polymorphisms in advanced colorectal cancer patients correlated with response to anti-EGFR antibodies and clinical outcome. *J. Transl. Med.* **10**, 232 (2012).
- 116. Etienne-Grimaldi, M.-C. *et al.* Multifactorial pharmacogenetic analysis in colorectal cancer patients receiving 5-fluorouracil-based therapy together with cetuximabirinotecan. *Br. J. Clin. Pharmacol.* **73**, 776–85 (2012).
- 117. Bibeau, F. *et al.* Impact of Fc{gamma}RIIa-Fc{gamma}RIIIa polymorphisms and KRAS mutations on the clinical outcome of patients with metastatic colorectal cancer treated with cetuximab plus irinotecan. *J. Clin. Oncol.* **27**, 1122–9 (2009).
- 118. Courtney, K. D., Corcoran, R. B. & Engelman, J. A. The PI3K pathway as drug target in human cancer. *J. Clin. Oncol.* **28**, 1075–83 (2010).
- 119. Katz, M., Amit, I. & Yarden, Y. Regulation of MAPKs by growth factors and receptor tyrosine kinases. *Biochim. Biophys. Acta* **1773**, 1161–76 (2007).
- Punt, C. J. A., Koopman, M. & Vermeulen, L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat. Rev. Clin. Oncol.* 14, 235– 246 (2017).
- 121. Pelizzo, M. R. *et al.* The role of BRAF(V600E) mutation as poor prognostic factor for the outcome of patients with intrathyroid papillary thyroid carcinoma. *Biomed. Pharmacother.* = *Biomédecine pharmacothérapie* **68**, 413–7 (2014).
- 122. Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N. Engl. J. Med.* **364**, 2507–16 (2011).
- 123. Agúndez, J. A. G., Blanca, M., Cornejo-García, J. A. & García-Martín, E. Pharmacogenomics of cyclooxygenases. *Pharmacogenomics* **16**, 501–22 (2015).
- 124. Papafili, A. *et al.* Common promoter variant in cyclooxygenase-2 represses gene expression: evidence of role in acute-phase inflammatory response. *Arterioscler. Thromb. Vasc. Biol.* **22**, 1631–6 (2002).
- 125. de Vries, N., Swets, M., Vahrmeijer, A., Hokland, M. & Kuppen, P. The Immunogenicity of Colorectal Cancer in Relation to Tumor Development and Treatment. *Int. J. Mol. Sci.* **17**, 1030 (2016).
- 126. Riley, J. M. *et al.* The clinical implications of immunogenomics in colorectal cancer: A path for precision medicine. *Cancer* (2018). doi:10.1002/cncr.31214
- Passardi, A., Canale, M., Valgiusti, M. & Ulivi, P. Immune checkpoints as a target for colorectal cancer treatment. *International Journal of Molecular Sciences* 18, (2017).

- 128. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer* **12**, 252–264 (2012).
- Wu, X. *et al.* PD-1(+) CD8(+) T cells are exhausted in tumours and functional in draining lymph nodes of colorectal cancer patients. *Br. J. Cancer* 111, 1391–1399 (2014).
- 130. Rosenbaum, M. W., Bledsoe, J. R., Morales-Oyarvide, V., Huynh, T. G. & Mino-Kenudson, M. PD-L1 expression in colorectal cancer is associated with microsatellite instability, BRAF mutation, medullary morphology and cytotoxic tumor-infiltrating lymphocytes. in *Modern Pathology* **29**, 1104–1112 (2016).
- 131. Wang, H. Bin *et al.* Rise of PD-L1 expression during metastasis of colorectal cancer: Implications for immunotherapy. *J. Dig. Dis.* **18**, 574–581 (2017).
- 132. Goswami, S., Basu, S. & Sharma, P. A potential biomarker for anti-PD-1 immunotherapy. *Nat. Med.* **24**, 123–124 (2018).
- 133. Sharpe, A. H. & Pauken, K. E. The diverse functions of the PD1 inhibitory pathway. *Nat. Rev. Immunol.* (2017). doi:10.1038/nri.2017.108
- Singh, P. P., Sharma, P. K., Krishnan, G. & Lockhart, A. C. Immune checkpoints and immunotherapy for colorectal cancer. *Gastroenterol. Rep.* gov053 (2015). doi:10.1093/gastro/gov053
- 135. Ribas, A. *et al.* Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): A randomised, controlled, phase 2 trial. *Lancet Oncol.* **16**, 908–918 (2015).
- 136. Brahmer, J. *et al.* Nivolumab versus Docetaxel in Advanced Squamous-Cell Non– Small-Cell Lung Cancer. *N. Engl. J. Med.* **373**, 123–135 (2015).
- 137. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
- 138. Lipson, E. J. *et al.* Durable cancer regression off-treatment and effective reinduction therapy with an anti-PD-1 antibody. *Clin. Cancer Res.* **19**, 462–468 (2013).
- Llosa, N. J. *et al.* The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov.* 5, 43–51 (2015).
- 140. Gatalica, Z. *et al.* Programmed cell death 1 (PD-1) and its ligand (PD-L1) in common cancers and their correlation with molecular cancer type. *Cancer Epidemiol. Biomarkers Prev.* **23**, 2965–2970 (2014).
- Tauriello, D. V. F. *et al.* TGF-beta drives immune evasion in genetically reconstituted colon cancer metastasis. *Nat. Publ. Gr.* (2018). doi:10.1038/nature25492

- 142. Cantelli, G., Crosas-Molist, E., Georgouli, M. & Sanz-Moreno, V. TGFB-induced transcription in cancer. *Semin. Cancer Biol.* **42**, 60–69 (2017).
- 143. Colak, S. & ten Dijke, P. Targeting TGF-β Signaling in Cancer. *Trends in Cancer* **3**, 56–71 (2017).
- 144. Ding, L. *et al.* Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. *Cell* **173**, 305–320.e10 (2018).
- Huang, K. lin *et al.* Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell* 355–370 (2018). doi:10.1016/j.cell.2018.03.039
- 146. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
- 147. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447-52 (2015).
- 148. Thorn, C. F., Klein, T. E. & Altman, R. B. PharmGKB: The pharmacogenomics knowledge base. *Methods Mol. Biol.* **1015**, 311–320 (2013).
- 149. Amstutz, U. *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Dihydropyrimidine Dehydrogenase Genotype and Fluoropyrimidine Dosing: 2017 Update. *Clin. Pharmacol. Ther.* **103**, 210–216 (2018).
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57 (2009).
- Saito, Y. *et al.* CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clinical Pharmacology and Therapeutics* 99, 36–37 (2016).
- 152. Swen, J. J. *et al.* Pharmacogenetics: From Bench to Byte— An Update of Guidelines. *Clin. Pharmacol. Ther.* **89**, 662–673 (2011).
- 153. Ross, C. J. D. *et al.* The Canadian Pharmacogenomics Network for Drug Safety: A Model for Safety Pharmacology. *Thyroid* **20**, 681–687 (2010).
- 154. Quinones, L. *et al.* Perception of the Usefulness of Drug/Gene Pairs and Barriers for Pharmacogenomics in Latin America. *Curr. Drug Metab.* **15**, 202–208 (2014).
- 155. Johnson, J. A. & Cavallari, L. H. Pharmacogenetics and Cardiovascular Disease--Implications for Personalized Medicine. *Pharmacol. Rev.* **65**, 987–1009 (2013).

### **FIGURE LEGENDS**

**Figure 1.** Integrating multi-omic features in CRC subtypes. Microsatellite instability (MSI) is linked to hypermutation, hypermethylation, highly immunogenic response, and locations in the proximal colon (consensus molecular subtype 1 (CMS1)). Tumors with chromosomal instability (CIN) are linked to copy number variations, poorly immunogenic or inflamed, non-hypermutated subtypes, stromal infiltration, and locations in left colon or rectum (CMS2, CMS3, and CMS4).

**Figure 2.** Overview of different drugs used in the CRC treatments: a) antiangiogenics, b) drugs against EGFR, c) antimetabolites, d) platinum derivatives, e) immune checkpoint inhibitors (anti PD-1, anti PD-L1 and TGF $\beta$  inhibitor), and f) antitopoisomerase.

**Figure 3.** Biomarker network in CRC made up of driver genes, genes with pathogenic germline mutations, genes with somatic mutations, and druggable enzymes by antitopoisomerase, antimetabolite, platinum derivative and antiangiogenic drugs. The PPi network with a highest confidence cutoff of 0.9 was created using String Database.

**Figure 4.** Minor allele frequencies of druggable enzymes studied in 8674 samples from Latin American populations, and its relation with the category of drugs applied in CRC treatments.

# **TABLES**

Category	Agents	Drugs
Cytotoxic agents	Platinum derivatives	Oxaliplatin
	Antimetabolites	5-FU/leucovorin, capecitabine
	Antitopoisomerase	Irinotecan
	Nucleoside analogue <sup>a</sup>	Trifluridine, tripiracil
Monoclonal antibodies	Antiangiogenic <sup>b</sup>	Bevacizumab, ziv-aflivercept, ramucirumab
	Anti EGFR <sup>b</sup>	Cetuximab, panitumumab
	Anti BRAF V600E <sup>b</sup>	Vemurafenib
	Anti PD-L1, PD-1 <sup>b</sup>	Nivolumab, pembrolizumab
Tyrosine kinase inhibitors <sup>a</sup>	Poli anti kinase inhibitors	Regorafenib, sorafenib

# Table 1. Category of drugs applied in CRC treatments

5-FU, 5-fluorouracil; EGFR, epidermal growth factor receptor; BRAF, B-Raf protooncogene; PD-1, programmed cell death protein 1; PD-L1, programmed cell death ligand 1. <sup>a</sup>For patients who have progressed through all available regimens. <sup>b</sup>For advanced or metastatic disease only.

## Table 2. Biomarkers focused on oxaliplatin

Gene	Polymorphism	Clinical relevance	Function	Type of inheritance	Reference
GSTP1	rs1695 (A313G)	Neurotoxicity, neutropenia	Enzyme	Germinal	50,52
GSTM1	Del	Poor survival, neutropenia	Enzyme	Germinal	52
ERCC1	rs11615 (T354C)	Neuropathy / Survival	Repair protein	Germinal	50
ERCC2	rs13181 (A2251C/T)	Effectiveness / Survival	Repair protein	Germinal	42,48
ERCC4	rs1799801 (T2505C)	Response	Repair protein	Germinal	53
ERCC5	rs2016073 (A-763G);	Response	Repair protein	Germinal	49
	rs751402 (A+25G)				
XRCC1	rs25487 (G1196A)	Response	Repair protein	Somatic	50,54
XRCC3	rs1799794 (A316G)	Neutropenia	Repair protein	Germinal	55

GSTP1, glutathione S-transferase pi 1; GSTM1, glutathione S-transferase mu 1; ERCC1, excision repair 1; ERCC2, excision repair 2; ERCC4, excision repair 4; ERCC5, excision repair 5; XRCC1, X-ray repair cross complementing 1; XRCC3, X-ray repair cross complementing 3; del, deletion; G, guanine; A, adenine; C, cytosine; T, thymine.

**Table 3.** Allele frequencies for clinically relevant genetic variants GSTP1 rs1695, ERCC1 rs11615, ERCC2 rs13181, ERCC4 rs1799801, ERCC5 rs2016073, CXCR1 rs2234671, XRCC1 rs25487 and XRCC3 rs1799794 in populations worldwide

Como	Dolymounhism		Human Populations	
Gene	Polymorphism	Latin American	Caucasian	Asian
		Colombia: 0.36 (G) <sup>a</sup> ;	Spain: 0.36 (G);	Han Chinese: 0.18 (G);
CCTD1	rs1695	Mexico: 0.56 (G);	British: 0.32 (G);	Bangladesh: 0.22 (G);
USIPI	(A313G)	Peru: 0.67 (G);	Finland: 0.28 (G);	Japan: 0.10 (G);
		Puerto Rico: 0.37 (G)	Italy: 0.29 (G)	Vietnam: 0.20 (G)
		Colombia: 0.52 (G);	Spain: 0.37 (G);	Han Chinese: 0.75 (G);
EDCC1	rs11615	Mexico: 0.74 (G);	British: 0.32 (G);	Bangladesh: 0.65 (G);
ERCCI	(T354C)	Peru: 0.75 (G);	Finland: 0.37 (G);	Japan: 0.71 (G);
		Puerto Rico: 0.50 (G)	Italy: 0.46 (G)	Vietnam: 0.73 (G)
		Colombia: 0.24 (G);	Spain: 0.31 (G);	Han Chinese: 0.11 (G);
EDCC2	rs13181	Mexico: 0.19 (G);	British: 0.30 (G);	Bangladesh: 0.35 (G);
ERCC2	(A2251C/T)	Peru: 0.17 (G);	Finland: 0.40 (G);	Japan: 0.07 (G);
		Puerto Rico: 0.24 (G)	Italy: 0.45 (G)	Vietnam: 0.08 (G)
		Colombia: 0.22 (G);	Spain: 0.38 (G);	Han Chinese: 0.21 (G);
EDCC4	rs1799801	Mexico: 0.19 (G);	British: 0.28 (G);	Bangladesh: 0.24 (G);
ERCC4	(T2505C)	Peru: 0.24 (G);	Finland: 0.25 (G);	Japan: 0.33 (G);
		Puerto Rico: 0.19 (G)	Italy: 0.29 (G)	Vietnam: 0.37 (G)
		Colombia: 0.73 (A);	Spain: 0.87 (A);	Han Chinese: 0.65 (A);
EDCC5	rs2016073	Mexico: 0.61 (A);	British: 0.80 (A);	Bangladesh: 0.66 (A);
ERCUS	(A-763G)	Peru: 0.48 (A);	Finland: 0.82 (A);	Japan: 0.73 (A);
		Puerto Rico: 0.78 (A)	Italy: 0.77 (A)	Vietnam: 0.63 (A)
		Colombia: 0.08 (G);	Spain: 0.01 (G);	Han Chinese: 0.10 (G);
CVCD1	rs2234671	Mexico: 0.14 (G);	British: 0.08 (G);	Bangladesh: 0.18 (G);
CACKI	(G2607C)	Peru: 0.31 (G);	Finland: 0.04 (G);	Japan: 0.08 (G);
		Puerto Rico: 0.08 (G)	Italy: 0.03 (G)	Vietnam: 0.06 (G)
		Colombia: 0.63 (C);	Spain: 0.58 (C);	Han Chinese: 0.75 (C);
VDCC1	rs25487	Mexico: 0.73 (C);	British: 0.66 (C);	Bangladesh: 0.66 (C);
ARCCI	(G1196A)	Peru: 0.69 (C);	Finland: 0.67 (C);	Japan: 0.72 (C);
		Puerto Rico: 0.71 (C)	Italy: 0.63 (C)	Vietnam: 0.77 (C)
		Colombia: 0.19 (C);	Spain: 0.28 (C);	Han Chinese: 0.52 (C);
VPCC2	rs1799794	Mexico: 0.16 (C);	British: 0.19 (C);	Bangladesh: 0.42 (C);
лкссэ	(A316G)	Peru: 0.25 (C);	Finland: 0.23 (C);	Japan: 0.42 (C);
		Puerto Rico: 0.16 (C)	Italy: 0.19 (C)	Vietnam: 0.39 (C)

GSTP1, glutathione S-transferase pi 1; ERCC1, excision repair 1; ERCC2, excision repair 2; ERCC4, excision repair 4; ERCC5, excision repair 5; XRCC1, X-ray repair cross complementing 1; XRCC3, X-ray repair cross complementing 3; CXCR1, C-X-C motif chemokine receptor 1; G, guanine; A, adenine; C, cytosine; T, thymine. <sup>a</sup> Frequency of minor allele.

Gene	Polymorphism	Clinical relevance	Function	Type of inheritance	Reference
TYMS	rs45445694 (*2R/*3R)	Neutropenia	Enzyme	Germinal	63,70
DPYD	rs3918290 (G1905+1A);	Toxicity	Enzyme	Germinal	63,64,71
	rs67376798 (A2846T);	Toxicity			
	rs1801158 (G1601A);	Toxicity			
	rs55886062 (T1679G);	Toxicity			
	rs1801159 (A1627G);	Neutropenia			
	rs12132152 (G97057448A);	Diarrhea			
	rs12022243 (C97397224T)	Diarrhea			
MTHFR	rs1801131 (A1298T);	Toxicity	Enzyme	Germinal	70,72
	rs1801133 (C677T)	Toxicity			
ABCB1	rs1128503 (C1236T);	Neutropenia	Transporter	Germinal	67,68
	rs1045642 (C3435T)	Diarrhea			
CDA	rs2072671 (A79C);	Toxicity	Enzyme	Germinal	67
	rs602950 (A-92G);	Diarrhea	-		
	rs532545 (C-451T)	Diarrhea			
CCND1	rs9344 (G870A)	Prognosis	Cyclin	Germinal	73
TP	rs11479 (C1412T)	Response	Enzyme	Germinal	69
ENOSF1	rs2612091 (805-227G>A)	Diarrhea	Enzyme	Germinal	64,67

Table 4. Biomarkers focused on capecitabine, 5-FU and gemcitabine drugs

ABCB1, ATP binding cassette subfamily B member 1; CCND1, cyclin D1; CDA, cytidine deaminase; DPYD, dihydropyrimidine dehydrogenase; ENOSF1, enolase superfamily member 1; MTHFR, methylenetetrahydrofolate reductase; TP, thymidine phosphorylase; TYMS, thymidylate synthetase; G, guanine; A, adenine; C, cytosine; T, thymine.

**Table 5.** Allele frequencies for clinically relevant genetic variants DPYD rs3918290, MTHFR rs1801133, ABCB1 rs1128503, CDA rs2072671, CCND1 rs9344, TP rs9344 and ENOSF1 rs2612091 in populations worldwide

Com	D. I		Human Populations	
Gene	Polymorphism	Latin American	Caucasian	Asian
		Colombia: 0.00 (T) <sup>a</sup> ;	Spain: 0.00 (T);	Han Chinese: 0.00 (T);
DBVD	rs3918290	Mexico: 0.00 (T);	British: 0.00 (T);	Bangladesh: 0.00 (T);
DFID	(1905+1G>A)	Peru: 0.01 (T);	Finland: 0.02 (T);	Japan: 0.00 (T);
		Puerto Rico: 0.00 (T)	Italy: 0.00 (T)	Vietnam: 0.00 (T)
		Colombia: 0.54 (A);	Spain: 0.44 (A);	Han Chinese: 0.47 (A);
MTHED	rs1801133	Mexico: 0.47 (A);	British: 0.32 (A);	Bangladesh: 0.12 (A);
МІПГК	(C677T)	Peru: 0.44 (A);	Finland: 0.27 (A);	Japan: 0.38 (A);
		Puerto Rico: 0.45 (A)	Italy: 0.47 (A)	Vietnam: 0.19 (A)
		Colombia: 0.57 (G);	Spain: 0.62 (G);	Han Chinese: 0.30 (G);
ADCD1	rs1128503	Mexico: 0.53 (G);	British: 0.58 (G);	Bangladesh: 0.37 (G);
ADCDI	(C1236T)	Peru: 0.67 (G);	Finland: 0.57 (G);	Japan: 0.40 (G);
		Puerto Rico: 0.60 (G)	Italy: 0.58 (G)	Vietnam: 0.42 (G)
		Colombia: 0.27 (C);	Spain: 0.34 (C);	Han Chinese: 0.12 (C);
CDA	rs2072671	Mexico: 0.32 (C);	British: 0.33 (C);	Bangladesh: 0.17 (C);
CDA	(A79C)	Peru: 0.36 (C);	Finland: 0.19 (C);	Japan: 0.21 (C);
		Puerto Rico: 0.28 (C)	Italy: 0.37 (C)	Vietnam: 0.10 (C)
		Colombia: 0.33 (A);	Spain: 0.55 (A);	Han Chinese: 0.56 (A);
CONDI	rs9344	Mexico: 0.33 (A);	British: 0.47 (A);	Bangladesh: 0.58 (A);
CUNDI	(G870A)	Peru: 0.30 (A);	Finland: 0.42 (A);	Japan: 0.47 (A);
		Puerto Rico: 0.42 (A)	Italy: 0.51 (A)	Vietnam: 0.63 (A)
		Colombia: 0.11 (A);	Spain: 0.06 (A);	Han Chinese: 0.24 (A);
TD	rs11479	Mexico: 0.25 (A);	British: 0.05 (A);	Bangladesh: 0.16 (A);
IP	(C1412T)	Peru: 0.24 (A);	Finland: 0.08 (A);	Japan: 0.25 (A);
		Puerto Rico: 0.12 (A)	Italy: 0.07 (A)	Vietnam: 0.32 (A)
		Colombia: 0.61 (T);	Spain: 0.54 (T);	Han Chinese: 0.70 (T);
ENOSEI	rs2612091	Mexico: 0.59 (T);	British: 0.53 (T);	Bangladesh: 0.55 (T);
ENUSFI	(805-227G>A)	Peru: 0.69 (T);	Finland: 0.57 (T);	Japan: 0.69 (T);
		Puerto Rico: 0.61 (T)	Italy: 0.56 (T)	Vietnam: 0.70 (T)

ABCB1, ATP binding cassette subfamily B member 1; CCND1, cyclin D1; CDA, cytidine deaminase; DPYD, dihydropyrimidine dehydrogenase; ENOSF1, enolase superfamily member 1; MTHFR, methylenetetrahydrofolate reductase; TP, thymidine phosphorylase; G, guanine; A, adenine; C, cytosine; T, thymine. <sup>a</sup> Frequency of minor allele.

Gene	Polymorphism	Clinical relevance	Function	Type of inheritance	Reference
CYP3A5	rs776746 (*3C)	Response	Enzyme	Germinal	45
UGT1A1	rs8175347 (*28);	Neutropenia and Diarrhea	Enzyme	Germinal	78,79,81
	rs4148323 (*6)	Neutropenia			
UGT1A7	rs17868324 (*3)	Neutropenia	Enzyme	Germinal	79,81
UGT1A9	rs3832043 (*22)	Neutropenia	Enzyme	Germinal	79,81,88
ABCB1	rs1128503 (C1236T);	Asthenia	Transporter	Germinal	67,68
	rs1045642 (C3435T);	Diarrhea	_		
	rs2032582 (G2677T/A)	Global Survival			
ABCC1	rs2074087 (C2461-30G)	Neutropenia	Transporter	Germinal	52,55
ABCC2	rs3740066 (T3972C)	Neutropenia	Transporter	Germinal	86,87,89
ABCG2	rs262604 (-20+805 A>G);	Myelosuppression	Transporter	Germinal	83
	rs2231142 (C421A);	Neutropenia	_		
	rs7699188 (C61414T)	Toxicity			
SLC19A1	rs1051266 (A80G)	PFS	Transporter	Germinal	84
SLCO1B1	rs2306283 (A388G)	PFS	Transporter	Germinal	84

## Table 6. Biomarkers focused on irinotecan drug

ABCB1, ATP binding cassette subfamily B member 1; ABCC1, ATP binding cassette subfamily C member 1; ABCC2, ATP binding cassette subfamily C member 2; ABCG2, ATP binding cassette subfamily G member 2; PFS, progression-free survival; CYP3A5, cytochrome P450 family 3 subfamily A member 5; UGT1A1, UDP glucuronosyltransferase family 1 member A1; UGT1A7, UDP glucuronosyltransferase family 1 member A7; UGT1A9, UDP glucuronosyltransferase family 1 member A9; SLC19A1, solute carrier family 19 member 1; SLCO1B1, solute carrier organic anion transporter family member 1B1; G, guanine; A, adenine; C, cytosine; T, thymine; \*, repetitions

**Table 7.** Allele frequencies for clinically relevant germline polymorphisms CES1 rs2244613, ABCB1 rs1045642, ABCC1 rs2074087, ABCG2 rs262604, SCL19A1 rs1051266 and SLC01B1 rs2306283 in populations worldwide

Como	Dolymounhism	Human Populations			
Gene	Polymorphism	Latin American	Caucasian	Asian	
		Colombia: 0.80 (T) <sup>a</sup> ;	Spain: 0.80 (T);	Han Chinese: 0.38 (T);	
CES1	rs2244613	Mexico: 0.70 (T);	British: 0.84 (T);	Bangladesh: 0.60 (T);	
CESI	(C1168-33A)	Peru: 0.64 (T);	Finland: 0.84 (T);	Japan: 0.37 (T);	
		Puerto Rico: 0.75 (T)	Italy: 0.87 (T)	Vietnam: 0.43 (T)	
		Colombia: 0.56 (G);	Spain: 0.54 (G);	Han Chinese: 0.62 (G);	
ADCD1	rs1045642	Mexico: 0.52 (G);	British: 0.47 (G);	Bangladesh: 0.39 (G);	
ADCDI	(C3435T)	Peru: 0.62 (G);	Finland: 0.42 (G);	Japan: 0.52 (G);	
		Puerto Rico: 0.57 (G)	Italy: 0.53 (G)	Vietnam: 0.60 (G)	
		Colombia: 0.79 (G);	Spain: 0.85 (G);	Han Chinese: 0.78 (G);	
APCCI	rs2074087	Mexico: 0.77 (G);	British: 0.88 (G);	Bangladesh: 0.57 (G);	
ABCCI	(C2461-30G)	Peru: 0.77 (G);	Finland: 0.84 (G);	Japan: 0.78 (G);	
		Puerto Rico: 0.78 (G)	Italy: 0.79 (G)	Vietnam: 0.88 (G)	
	rs3740066 (T3972C)	Colombia: 0.36 (T);	Spain: 0.39 (T);	Han Chinese: 0.25 (T);	
ABCC2		Mexico: 0.42 (T);	British: 0.38 (T);	Bangladesh: 0.33 (T);	
ABCC2		Peru: 0.24 (T);	Finland: 0.37 (T);	Japan: 0.23 (T);	
		Puerto Rico: 0.36 (T)	Italy: 0.37 (T)	Vietnam: 0.27 (T)	
		Colombia: 1.00 (C);	Spain: 1.00 (C);	Han Chinese: 1.00 (C);	
APCC2	rs262604	Mexico: 1.00 (C);	British: 1.00 (C);	Bangladesh: 1.00 (C);	
ABC02	(-20+805A>G)	Peru: 1.00 (C);	Finland: 1.00 (C);	Japan: 1.00 (C);	
		Puerto Rico: 1.00 (C)	Italy: 1.00 (C)	Vietnam: 1.00 (C)	
		Colombia: 0.51 (G);	Spain: 0.49 (G);	Han Chinese: 0.52 (G);	
SI C10A1	rs1051266	Mexico: 0.65 (G);	British: 0.60 (G);	Bangladesh: 0.62 (G);	
SLCIFAI	(A80G)	Peru: 0.63 (G);	Finland: 0.55 (G);	Japan: 0.46 (G);	
		Puerto Rico: 0.57 (G)	Italy: 0.55 (G)	Vietnam: 0.46 (G)	
		Colombia: 0.48 (G);	Spain: 0.42 (G);	Han Chinese: 0.78 (G);	
SI CO1P1	rs2306283	Mexico: 0.38 (G);	British: 0.36 (G);	Bangladesh: 0.56 (G);	
SLCOIDI	(A388G)	Peru: 0.47 (G);	Finland: 0.44 (G);	Japan: 0.66 (G);	
		Puerto Rico: 0.53 (G)	Italy: 0.39 (G)	Vietnam: 0.78 (G)	

CES1, carboxylesterase 1; ABCB1, ATP binding cassette subfamily B member 1; ABCC1, ATP binding cassette subfamily C member 1; ABCC2, ATP binding cassette subfamily C member 2; ABCG2, ATP binding cassette subfamily G member 2; SLC19A1, solute carrier family 19 member 1; SLCO1B1, solute carrier organic anion transporter family member 1B1; G, guanine; A, adenine; C, cytosine; T, thymine. <sup>a</sup> Frequency of minor allele.

1 million 1					÷
Gene	Polymorphism	Clinical relevance	Function	Type of inheritance	Reference
VEGFR1	rs9582036 (C-834A)	OS	Receptor	Germinal	94
VEGFR2	rs12505758 (2266+1166 A>G)	PFS	Receptor	Germinal	92,94
VEGFA	rs3025039 (C*237T)	PFS	Growth factor	Germinal	94
	rs13207351 (G-152A)	PFS		Germinal	
ANXA11	rs1049550 (C688T)	ORR	Calcium-dependent	Germinal	95
			phospholipid-binding		
			proteins		
CXCR1	rs2234671 (G827C)	Response	Receptor	Germinal	93
CXCR2	rs2230054 (C786T)	ORR	Receptor	Germinal	93

### Table 8. Biomarkers focused on antiangiogenics and bevacizumab drug

ANXA11, annexin A11; CXCR1, C-X-C motif chemokine receptor 1; CXCR2, C-X-C motif chemokine receptor 2; VEGFA, vascular endothelial growth factor A; VEGFR1, vascular endothelial growth factor receptor 1; VEGFR2, vascular endothelial growth factor receptor 2; G, guanine; A, adenine; C, cytosine; T, thymine; OS, overall survival; PFS, patient-free survival; ORR, overall response rates.

**Table 9.** Allele frequencies for clinically relevant germline polymorphisms VEGFR1 rs9582036, VEGFR2 rs12505758, VEGFA rs3025039, ANXA11 rs1049550 and CXCR2 rs2230054 in populations worldwide

Como	Dalumannhiam	Human Populations			
Gene	Polymorphism	Latin American	Caucasian	Asian	
		Colombia: 0.73 (A) <sup>a</sup> ;	Spain: 0.69 (A);	Han Chinese: 0.83 (A);	
VECED1	rs9582036	Mexico: 0.84 (A);	British: 0.70 (A);	Bangladesh: 0.84 (A);	
VEGFR1	(C-834A)	Peru: 0.94 (A);	Finland: 0.78 (A);	Japan: 0.85 (A);	
		Puerto Rico: 0.64 (A)	Italy: 0.75 (A)	Vietnam: 0.79 (A)	
		Colombia: 0.13 (C);	Spain: 0.12 (C);	Han Chinese: 0.19 (C);	
VECEDO	rs12505758	Mexico: 0.17 (C);	British: 0.07 (C);	Bangladesh: 0.34 (C);	
VEGFR2	(2266+1166 A>G)	Peru: 0.31 (C);	Finland: 0.10 (C);	Japan: 0.27 (C);	
		Puerto Rico: 0.09 (C)	Italy: 0.12 (C)	Vietnam: 0.15 (C)	
		Colombia: 0.13 (T);	Spain: 0.13 (T);	Han Chinese: 0.18 (T);	
VECEA	rs3025039	Mexico: 0.30 (T);	British: 0.07 (T);	Bangladesh: 0.12 (T);	
VEGFA	(C*237T)	Peru: 0.34 (T);	Finland: 0.14 (T);	Japan: 0.16 (T);	
		Puerto Rico: 0.18 (T)	Italy: 0.12 (T)	Vietnam: 0.16 (T)	
		Colombia: 0.45 (A);	Spain: 0.47 (A);	Han Chinese: 0.63 (A);	
A N 187 A 1 1	rs1049550	Mexico: 0.39 (A);	British: 0.44 (A);	Bangladesh: 0.35 (A);	
ANAATI	(C688T)	Peru: 0.64 (A);	Finland: 0.54 (A);	Japan: 0.65 (A);	
		Puerto Rico: 0.38 (A)	Italy: 0.41 (A)	Vietnam: 0.59 (A)	
		Colombia: 0.45 (T);	Spain: 0.51 (T);	Han Chinese: 0.36 (T);	
CYCD2	rs2230054	Mexico: 0.45 (T);	British: 0.47 (T);	Bangladesh: 0.49 (T);	
CACR2	(C786T)	Peru: 0.57 (T);	Finland: 0.42 (T);	Japan: 0.28 (T);	
		Puerto Rico: 0.54 (T)	Italy: 0.50 (T)	Vietnam: 0.39 (T)	

ANXA11, annexin A11; CXCR2, C-X-C motif chemokine receptor 2; VEGFA, vascular endothelial growth factor A; VEGFR1, vascular endothelial growth factor receptor 1; VEGFR2, vascular endothelial growth factor receptor 2; G, guanine; A, adenine; C, cytosine; T, thymine. <sup>a</sup> Frequency of minor allele.

Gene	Polymorphism	Clinical relevance	Function	Type of inheritance	Reference
EGFR	rs2227983 (G1562A);	Better prognosis	Receptor	Somatic	34,102-104
	rs712830 (C-191A);	Toxicity		Germinal	
	rs1050171 (G2226A);	PFS		Somatic	
	CA-repeat in intron 1	PFS		Germinal	
	rs1057519860 (S492R, A1474C)	Response		Somatic	
	S464L	Treatment resistance		Somatic	
	G465R	Treatment resistance		Somatic	
	I491M	Treatment resistance		Somatic	
	rs377567759 (R451C, C1351T)	Treatment resistance		Somatic	
	K467T	Treatment resistance		Somatic	
EGF	rs4444903 (G61A)	Response	Growth factor	Germinal	105
FCGR2A	rs1801274 (A535G)	Survival	Receptor	Germinal	62
FCGR3A	rs396991 (A818C)	Survival	Receptor	Germinal	102
PIK3CA	Exon 9 and 20 (Mutations)	Treatment resistance	Oncogene	Somatic	106
	consider alternative therapy for				
	cetuximab and panitumumab				
B-RAF	V600E: if the mutation is present	Treatment resistance	Protein kinase	Somatic	107,108
	considering alternative therapy				
	for cetuximab and panitumumab				
KRAS	Exon 2 (codons 12 and 13) and	Treatment resistance	Protooncogene	Somatic	100,101
	exon 4 (codon 61). If the		c		
	mutation is present, not use				
	cetuximab or panitumumab				
NRAS	Exon 2 (codons 12 and 13), exon	Treatment resistance	Protooncogene	Somatic	100,101
	3 (59 and 61), and exon 4 (117		-		
	and 146)				
COX2	rs20417 (G-765C)	PES	Enzyme	Germinal	63

**Table 10.** Biomarkers focused on gefitinib, erlotinib, imatinib, vemurafenib, cetuximab, and panitumumab drugs

EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; FCGR2A, Fc fragment of IgG receptor IIa; FCGR3A, Fc fragment of IgG receptor IIIa; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha; KRAS, KRAS proto-oncogene, GTPase; NRAS, neuroblastoma RAS viral oncogene homolog; COX2, cytochrome c oxidase subunit II; PFS, progression-free survival; G, guanine; A, adenine; C, cytosine; T, thymine.

**Table 11.** Allele frequencies for clinically relevant genetic variants EGFR rs2227983, EGF rs4444903, FGGR2A rs1801274, FCGR3A rs396991 and COX2 rs20417 in populations worldwide

Como	Delymouphism	Human Populations			
Gene	Polymorphism	Latin American	Caucasian	Asian	
		Colombia: 0.35 (A) <sup>a</sup> ;	Spain: 0.25 (A);	Han Chinese: 0.46 (A);	
EGFR	rs2227983	Mexico: 0.31 (A);	British: 0.24 (A);	Bangladesh: 0.34 (A);	
	(G1562A)	Peru: 0.36 (A);	Finland: 0.38 (A);	Japan: 0.62 (A);	
		Puerto Rico: 0.30 (A)	Italy: 0.25 (A)	Vietnam: 0.53 (A)	
		Colombia: 0.51 (G);	Spain: 0.39 (G);	Han Chinese: 0.70 (G);	
ECE	rs4444903	Mexico: 0.62 (G);	British: 0.41 (G);	Bangladesh: 0.64 (G);	
EGF	(G61A)	Peru: 0.70 (G);	Finland: 0.37 (G);	Japan: 0.71 (G);	
		Puerto Rico: 0.52 (G)	Italy: 0.38 (G)	Vietnam: 0.70 (G)	
		Colombia: 0.39 (G);	Spain: 0.53 (G);	Han Chinese: 0.34 (G);	
ECCD2A	rs1801274	Mexico: 0.51 (G);	British: 0.61 (G);	Bangladesh: 0.36 (G);	
FCGK2A	(A535G)	Peru: 0.47 (G);	Finland: 0.54 (G);	Japan: 0.19 (G);	
		Puerto Rico: 0.45 (G)	Italy: 0.41 (G)	Vietnam: 0.28 (G)	
		Colombia: 0.00 (C);	Spain: 0.00 (C);	Han Chinese: 0.00 (C);	
ECCD2A	rs396991	Mexico: 0.00 (C);	British: 0.00 (C);	Bangladesh: 0.00 (C);	
FUGKJA	(A818C)	Peru: 0.00 (C);	Finland: 0.00 (C);	Japan: 0.00 (C);	
		Puerto Rico: 0.00 (C)	Italy: 0.00 (C)	Vietnam: 0.01 (C)	
		Colombia: 0.22 (G);	Spain: 0.15 (G);	Han Chinese: 0.05 (G);	
COV2	rs20417	Mexico: 0.21 (G);	British: 0.14 (G);	Bangladesh: 0.17 (G);	
COA2	(G-765C)	Peru: 0.21 (G);	Finland: 0.11 (G);	Japan: 0.04 (G);	
		Puerto Rico: 0.21 (G)	Italy: 0.19 (G)	Vietnam: 0.02 (G)	

EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; FCGR2A, Fc fragment of IgG receptor IIa; FCGR3A, Fc fragment of IgG receptor IIIa; COX2, cytochrome c oxidase subunit II; G, guanine; A, adenine; C, cytosine; T, thymine. <sup>a</sup> Frequency of minor allele.

**Table 12.** Pathogenic germline variants in CRC according to the Pan-Cancer Atlas and allele frequencies according to the Exome Aggregation Consortium

Gene	Polymorphism	Alleles	Consequence	Variation	Overall	Frequency <sup>a</sup>
				type	classification	
APC	rs752519066 (L69*)	T>A	Stop gained	SNV	Likely pathogenic	A=0.00001
ATM	rs587782652 (V2716A)	T>C	Missense	SNV	Pathogenic	C=0.00004
ATR	rs777982083 (E409*)	C>T	Missense	SNV	Likely pathogenic	T=0.00002
	rs755272769	C>T	Splice	SNV	Likely pathogenic	T=0.00001
	(C142553741T)		acceptor			
	rs781260235 (L2093X)	delAG	Frameshift	Deletion	Likely pathogenic	delAG=0.00001
BARD1	rs587780021 (Q564*)	G>A	Stop gained	SNV	Pathogenic	A=0.00005
BLM	rs200389141 (Q548*)	C>T	Stop gained	SNV	Pathogenic	T=0.00018
BRCA1	rs80357669 (S819X)	delG	Frameshift	Deletion	Pathogenic	delG=0.00002
BRCA2	rs80359550 (S1982X)	delT	Frameshift	Deletion	Pathogenic	delT=0.00027
	rs80359013 (W2626C)	G>A, G>C	Missense	SNV	Pathogenic	C=0.00002
BRIP1	rs137852986 (R798*)	G>A	Stop gained	SNV	Pathogenic	A=0.00015
CHEK2	rs137853011 (S571F)	G>A	Missense	SNV	Pathogenic	A=0.00031
COL7A1	rs753819164 (R226*)	G>A	Stop gained	SNV	Likely pathogenic	A=0.00001
FANCI	rs121918164 (R1285*)	C>T	Stop gained	SNV	Pathogenic	T=0.00005
GJB2	rs766975999 (S222*)	G>T	Stop gained	SNV	Likely pathogenic	T=0.00002
MLH1	rs63751615 (R226*)	C>T	Stop gained	SNV	Pathogenic	T=0.00001
	rs780956158 (I691IX)	dupT	Frameshift	Insertion	Likely pathogenic	dupT=0.00001
MSH2	rs63749932 (R680*)	C>G, C>T	Stop gained	SNV	Pathogenic	G=0.0000
	rs760228651 (L407LX)	dupC	Frameshift	Insertion	Likely pathogenic	dupC=0.00001
MSH6	rs587781691 (R248X)	delC	Frameshift	Deletion	Pathogenic	delC=0.00001
	rs771764652 (SK536-537X)	delAGTA	Frameshift	Deletion	Likely pathogenic	delAGTA=0.00001
PALB2	rs515726124 (R170X)	-	Frameshift	Deletion	Pathogenic	-
	rs756660214 (L253LX)	-	Frameshift	Insertion	Pathogenic	-
POT1	rs750470470 (-357-358X)	dupA	Frameshift	Insertion	Likely pathogenic	dupA=0.00006
RAD51D	rs775045445 (W36*)	C>T	Stop gained	SNV	Likely pathogenic	T=0.00006
RECQL4	rs386833845 (C525X)	delA	Frameshift	Insertion	Pathogenic	delA=0.00026
RET	rs78347871 (R912P)	G>A, G>C,	Missense	SNV	Pathogenic	T=0.0000
		G>T				
RHBDF2	rs777871789 (W574*)	C>T	Stop gained	SNV	Likely pathogenic	T=0.00001
SDHA	rs766667009 (G251104T)	G>T	Splice donor	SNV	Pathogenic	T=0.00001

SNV, single nucleotide variant; \*, stop gained; dup, duplication; del, deletion; APC, APC WNT signaling pathway regulator; ATM, ATM serine/threonine kinase; ATR, ATR serine/threonine kinase; BARD1, BRCA1 associated RING domain 1; BLM, bloom syndrome RecQ like helicase; BRCA1, BRCA1 DNA repair associated; BRCA2, BRCA2 DNA repair associated; BRIP1, BRCA1 interacting protein C-terminal helicase 1; CHEK2, checkpoint kinase 2; COL7A1, collagen type VII alpha 1 chain; FANCI, Fanconi anemia complementation group I; GJB2, gap junction protein beta 2; MLH1, mutL homolog 1; MSH2, mutS homolog 2; MSH6, mutS homolog 6; PALB2, partner and localizer of BRCA2; POT1, protection of telomeres 1; RAD51D, RAD51 paralog D; RECQL4, RecQ like helicase 4; RET, ret proto-oncogene; RHBDF2, rhomboid 5 homolog 2; SDHA, succinate dehydrogenase complex flavoprotein subunit A. <sup>a</sup> Frequency of minor allele.

CMS1 immune	CMS3 metabolic	CMS2 canonical	CMS4 mesenchymal	
MSI Mutation count BRAF, MSH6, RNF43, ATM, TGFBR2 and PTEN mutations	KRAS and PIK3CA mutations	CIN Cop TP53, APC, KRAS, SMAD4 and PIK3CA mutations	y number variations APC and TP53 mutations	Genomics
Methylation RAS and BRAF mutations Proximal Tumor location Distal				Epigenomics Driver genes Clinical
Cancer-associated fibroblasts Highly immunogenic Poorly immunogenic Inflamed (immune-tolerant) Innate Immune response Adaptative				Stroma-immune microenvironment
High CpG islands Immune pathway expression JAK-STAT activation Caspases Right-side tumors Female	Low CpG islands Low WNT/MYC pathway activation IGFBPR2 overexpression Metabolic desregulation DNA damage repair	WNT/MYC pathway activation EGFR or SRC activation VEGF or VEGFR activation Integrins activation Better survival Left-side tumors	Stromal infiltration Immunosuppression TGFβ activation and EMT Angiogenesis NOTCH/EGFR2 overexpression Worse relapse-free over survival	Transcriptomic pathways and other facts

More (+)

Less (-)





CRC driver genes, genes with pathogenic germline mutations according to the Pan-Cancer Atlas, and genes with somatic mutations according to the Cancer Genome Interpreter

