Genetic Variations of Selected Genes Using Target Deep Sequencing in Colorectal Cancer Patients

Eman E Farghal1,2#, Marwa H Saied3#, Fatma M Ghaith1, Gamal I Moussa2,7, Gehan El-Sharnobi1, Samah M Soliman1,2, Hesham Tawfik6, Omnia Abdelfattah1, Enas A Abdul-Baki4, Fatma Gharib1, Lamis Mohamed6, Mohamed R El-Shanshory5,8, Amira Y Abdelnaby2,3, Mohamed L Salem2,4, Mona Watany1,2, Abdel Aziz A Zidan1, Yahia S Abdou1 and Said M Abdou1,2**

1Department of Clinical Pathology, Faculty of Medicine, Tanta University, Tanta, Egypt
2Genomic Signature Cancer Centre, Next Generation Sequencer Unit, Tanta University Global Educational Hospital, Tanta University, Tanta, Egypt
3Department of Clinical Pathology, Faculty of Medicine, Alexandria University, Alexandria, Egypt
4Department of Zoology, Division of Immunology and Biotechnology, Faculty of Science, Tanta University, Tanta, Egypt
5Department of Tropical Medicine, Faculty of Medicine, Tanta University, Tanta, Egypt
6Department of Clinical Oncology, Faculty of Medicine, Tanta University, Tanta, Egypt
7Department of Paediatrics, Faculty of Medicine, Tanta University, Tanta, Egypt
8Department of Zoology, Faculty of Science, Damanhur University, Damanhur, Egypt
9Department of Medicine, Ain Shams University, Egypt
*First and second authors equally contributed in this work

Abstract

Background: Colorectal carcinoma (CRC) is a burden problem in a developing country like Egypt since patients are usually admitted in late stage with bad prognosis and short overall survival. Because of genetic predisposition of CRC and introduction of advanced molecular techniques, efforts are directed to screen for potential pathogenic or disease-causing variants in CRC patients

Methods: DNA was isolated from formalin fixed paraffin embedded tissue sections collected from 24 CRC confirmed diagnosed patients. TruSight CRC panel (Illumina) was used for detection of different variants in 15 genes. The generated reads were obtained from Illumina Miseq were clustered into single nucleotide polymorphism (SNPs) and small insertions/deletions (Indels). Further pathogenic variants with somatic and germline mutations were identified according to the recommended criteria. Some CRC patients were subjected to anti-EGFR target therapy.

Results: Most of the variants were detected in TP53 gene 140 variants (65%); 105 short deletions none of them was pathogenic, 29 missense mutations and 6 SNPs at splicing sites. Next, ERBB2 has got 17 variants (8.8%) (missense and splicing), 8 of them were damaging disease causing variants. Besides, 16 pathogenic variants were identified in 12 patients (6 in TP53 and 7 in KRAS). Some pathogenic variants were not reported before in CRC e.g. TP53 C>A, rs121912654, Val157Phe. Additionally, patients carried different KRAS wild mutations showed variable response to anti-EGFR target therapy.

Conclusion: The most affected pathway in CRC was TP53 pathway followed by ERBB2, NRAS, KRAS and PIK3CA genes. Variable response to target therapy suggested dependence on the type of pathogenic variant identified, also a possible role of ERBB2 which had a significant variant frequency.

Keywords: Colorectal carcinoma; Trusight next generation sequencing; Molecular diagnosis

Introduction

Colorectal cancer (CRC) is considered the third most common tumour and the fourth leading cause of cancer-related mortality worldwide. Its incidence will be increased by 60% by the year 2030 [1]. At initial diagnosis, 25% of CRC patients present with metastasis and 50% will develop metastasis after diagnosis [2]. Many factors contributed to the development of CRC including genetic factors [3] and environmental factors such as diet and lifestyle [4].

Genetic alteration is a common feature of CRC ranging from small molecular changes such as point mutations or small indels to chromosomal copy number variation or translocation [5]. RAS status was mandatory to be examined before the use of anti-EGFR therapy as recommended by ESOM guidelines 2106. Also, patients carry BRAF mutations might not respond to anti-EGFR therapy [6]. Beside the RAS/RAF, MAP kinase pathway, TP53 and the PI3K/Akt/mTOR signalling cascade are critical. The process of CRC development is an interplay between suppression of some tumor suppressor genes (TSGs) including the key TSG i.e. TP53 and activation of some oncogenes such as KRAS, NRAS. The prevalence rate of TP53 mutation in Arab population is 52.5% compared to 47.5% in matched Western population [7]. TP53 mutations have roles in determining progression, invasiveness and also metastasis of CRC. So, CRC patients with mutant TP53 have more aggressive phenotype and poorer survival than those with TP53 wild type [8]. The phosphatidylinositol 3-kinase/Akt/mammalian target of
rapamycin (PI3K/Akt/mTOR) signalling pathway has pivotal roles in cell proliferation, apoptosis, survival and metastasis [9]. The effects of PI3K are mediated by AKT. One of the downstream targets of AKT is mTOR which controls angiogenesis. Activation of PI3K/Akt/mTOR has been reported in the development and progression of CRC [10].

Next-generation sequencing (NGS) technology has the advantage of parallel deep sequencing of genetic mutations in multiple genes in patients simultaneously [11]. Targeted sequencing is a good clinical application of NGS technology since it directed to increase the coverage of read depth of hotspot mutations through localizing the number of genes of interest examined with maintaining the number of bases sequenced [12]. Therefore, we aimed in this study to screen for potential pathogenic variants in Egyptian CRC patients in different stages of the disease.

Materials and Methods

Formalin fixed paraffin embedded tissues (FFPE) samples that were obtained from 24 CRC patients, who were referred to Tanta Main University hospital between January 2016 and June 2017. This study was approved by Tanta Ethics Committee of Faculty of Medicine. Informed consents were obtained from all patients. The age distribution of all patients ranged from 30 to 75 years with a median age of 55 years. The clinicopathological data of patients are summarized in Table 1. DNA was extracted from FFPE tissue samples using QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany). The concentration of DNA was determined using the Denovix fluorimeter (AGBL, USA).

Trusight tumor 15 panel illumina CRC panel

We sequenced patient’s DNA using the Trusight Tumor 15 panel (Illumina, San Diego, CA, USA) that contains 15 genes suspected to have roles in cancer predisposition; TP53, AKT1, BRAF, EGFR, ERBB2, FOX1L2, GNA11, GNAQ, KIT, KRAS, MET, NRAS, PDGFRα, PIK3CA and RET. Twelve genes showed coverage of 95% (≥ 20 reads). In these genes, the mean read coverage was 93.5% of bases covered at ≥ 500 × exonic and essential splice sites regions covered at ≥ 20 reads. Only, NRAS, RET and GNA11 showed coverage less than 95% (≤ 20 reads). For library preparation, Trusight panel DNA library preparation protocol (Illumina) was used [13].

Bioinformatic Analysis

The FASTQ files were generated using MiSeq Reporter (Illumina). The FASTQ files were assessed for their base quality. More than 85% of reads showed base quality higher than 30. Next, reads were aligned to a reference genome hg19 and variants were called using Genome Analysis Toolkit (GATK), Then, VCFs were generated for further analysis by VariantStudio (Illumina). The passed filtered variants were selected for the analysis. Annotations included in VariantStudio were HGMD, ClinVar pathogenicity annotations, links to dbSNP, in addition to prediction tools such as SIFT, PolyPhen and Provean. We also used Varsome (The Human Genomic Variant Search Engine, https://varsome.com/), Genome Aggregation Database/exomes (gnomAD, ExAC) and Mutation Taster. The identified variants in this study were classified according to ACMG (American Society of Medical Genetics and Genomics) into five categories: (1) pathogenic, (2) likely pathogenic, (3) of uncertain significance (VUS), (4) likely benign and (5) benign [14].

Results

Pathogenic and likely pathogenic variants in the CRC cohort

Sixteen pathogenic and likely pathogenic variants were identified in 12 patients, which account for 50% of patients (Table 2). All of those variants were single nucleotide polymorphism (SNPs); 7 SNPs in TP53 gene, 6 SNPs in KRAS, 1 SNP in BRAF and 2 SNPs in PIK3CA (Table 2).

TP53 pathogenic and likely pathogenic variants: Seven TP53 pathogenic/likely pathogenic variants were identified in the studied patients. There were no repeated TP53 mutations. Five out of seven TP53 mutations were germline inherited. All of these mutations were missense mutations located in three exons; 5, 7 and 8 (the commonest sites of TP53 mutations) [15]. Each exon has 2 identified mutations with one mutation found in exon 10. A pathogenic variant identified in patient ID 8 in exon 5 (C>A, rs121912654, Val157Phe), was not reported before in CRC, while it was found previously in patients with chronic lymphocytic leukemia [16] and hepatocellular carcinoma [17]. Similarly, a pathogenic variant was detected in exon 5 in patient ID 17 (C>T, rs28934578, Arg175His), is one of germline mutations detected in DNA binding domain in TP53 in solid tumors e.g. breast cancer [18]. Patient ID 17 also has got another pathogenic mutation (G>A, rs587782529, Arg337Cys) outside the common sites of TP53 mutations i.e. in exon 10. This mutation is associated with unclassical Li-Fraumeni syndrome [19]. Furthermore, exon 7 has got 2 mutations e.g. C>T, rs28934575, Gly245Ser which was previously described in CRC [20] and G>A, rs121912651, Arg248Trp which is a well-known pathogenic variant in CRC [21]. Exon 8 has got 2 pathogenic variants that are associated with Li-Fraumeni syndrome (G>A, rs149633775-Arg283Cys, C>T and rs763089116- Cys277Phe) [22]. Generally, TP53 has got 29 missense mutations; 7 were reported as pathogenic, 3 with uncertain significance and the remaining missense mutations were benign/likely benign.
KRAS pathogenic and likely pathogenic variants: Six pathogenic variants/likely pathogenic variants were identified in KRAS, 4 mutations were in exon 2 (66%), one mutation in exon 3 and one mutation in exon 4 (Table 2). All of them were SNPs and somatically inherited. KRAS exon 2 mutations are found to be predictors of bad prognosis and resistance to anti-EGFR antibody therapy [23]. Pathogenic variant (C>T, 112445441, Gly13Asp) was repeated in 2 patients; IDs 5 and 16. This mutation (Gly13Asp) was associated with invasive pancreaticobiliary tumors in Turkish population [24]. Pathogenic variant (rs121913529) was also identified in 2 patients; IDs 4 and 11. This mutation, in addition to the pathogenic variant (C>T, rs121913530, Gly12Ser) were detected in patient ID 8. Both variants were shown to discriminate conventional adenoma from CRC [25]. The remaining pathogenic KRAS variants were found in exon 3 (G>A, rs104894364, Thr58Ile patient ID 20) and in exon 4 (C>T, rs121913527, Ala146Thr patient ID 10), both are well-known pathogenic variants in CRC [26]. Furthermore, one pathogenic variant (A>T, rs104886003, Glu596Val) was identified in BRAF in patient ID 7, this mutation was detected in sessile serrated polyps; precursor of CRC [26] (Table 2). Also, Patient ID 19 has got C>A, Gly596Val variant. Finally, two pathogenic variants were detected in PIK3CA; G>A rs104886003, Glu545Lys in patient ID 4, this mutation is one of the hotspots found in exon 10 that is considered a driver PIK3CA gene mutation [27]. Also, a known pathogenic PIK3CA mutation (A>G, rs121913279, His1047Arg) [21] was identified in patient ID 16. Therefore, most of pathogenic missense mutations were identified in KRAS in exon 2. Less commonly, few hotspot pathogenic variants were detected in BRAF and PIK3CA.

Missense, splicing and indels mutations

Next, total number of mutations in every gene was investigated. TP53 has got the highest number of detected mutations in 24 patients i.e. 140 (65%); 105 short deletions none of them was pathogenic, 29 missense mutations and 6 SNPs at splicing sites, one of them was a disease-causing at intron 9 patient ID 22. ERBB2 was the second most common gene that had mutations (8.8%). ERBB2 has got 11 missense mutations; one in exon 17 (Arg678Trp), patient ID 8, four variants in exon 18; (Thr686Met) patient ID 7, (Glu717Asp) patient ID 14, (Ser728Phe) patient ID 3 and (Gly732Asp) patient ID 22. All of them were predicted as damaging by SIFT and Provean. However, three neutral tolerated missense variants in exon 20 were also identified in ERBB2: Val777Leu, Gly778Asp, Val777Leu and two tolerated neutral variants in exon 26 (Pro1121Leu, p.Val1128Ile). One missense variant predicted as damaging (Ser1054Ile) was identified in exon 26 in 3 patients IDs 4, 7 and 8. In addition, ERBB2 has got 6 variants at splice site; one damaging variant (Pro699Ser) in exon 18 patient ID 1, one damaging variant (Pro137Ala) in exon 26 patient ID 4. In addition, NRAS has got 6% of total identified variants. All ten missense mutations found in NRAS gene were of low frequency (seen in less than 10% of the reads covering a given base). Three low variant frequency mutations at the splice site were also identified; one in exon 3, two in exon 4; Gln99His patient ID 22 and Ser145Ter patient ID 19. Besides, KIT has got 3.8% of frequent detected variants; four missense mutations, three of them were predicted as damaging by SIFT and Provean; Gly432Val, Pro467Gln, Ser821Tyr, while a benign missense variant (Met541Leu) was identified. The latter mutation was repeated in patients IDs 18, 22 and 2. In addition, KIT included four mutations at the splice site, one of them was predicted as damaging by SIFT and Provean in exon 17 (Asn828Lys) in patient ID 22. One variant at the splice site (exon 18, G>C, rs3733542, benign, germline) was repeated in patients IDs 1, 15, 16, 18, 21, 22 and 24. PDGFRα has also 3.8% of the identified variants; 6 missense mutations (Met578Ile, Asn656Lys, Gly838Asp, Ala840Thr, Asn659Lys, and Ala663Asp). Besides, PDGFRα has a repeated inactivated silent SNP (A>G, rs1873778, at codon 567 (P567F) in exon 12 [28] that presents as homozygous allele (GG) in all CRC patients except in patients IDs 6, 20, 22 where the variant was heterozygous (AG). The GG allele was the most frequent similar to European Caucasian population [29]. A characteristic insertion (T>TA) in exon 18 was also found in most of patients (homozygous TA/TA in patients IDs 1, 2, 4, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 22, 23 and 24), (heterozygous T>TA, rs3830355, IVS18-50insA in patients IDs 3, 5, 6, 12 and 18). Both PDGFRα exon 18 mutation and exon 12 mutation were related more to gastric and intestinal GIST (gastrointestinal tumors) more than to colonic [30]. Two splice site variants were also identified without provided amino acid substitution. Similarly, KRAS has got 8 identified mutations (3.8%); six of them were pathogenic as mentioned before and 2 intronic SNPs. FOXL2 has 4 missense mutations Arg148Leu,  

<table>
<thead>
<tr>
<th>Gene</th>
<th>Case no.</th>
<th>Exon no</th>
<th>Type of mutation</th>
<th>Variant</th>
<th>Heterogeneity</th>
<th>Inheritance</th>
<th>SIFT</th>
<th>Provean</th>
<th>Amino acid substitution</th>
<th>Age at diagnosis (Years)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP53</td>
<td>17</td>
<td>Exon 10</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>p.Arg337Cys</td>
<td>75</td>
<td>rs587762529</td>
</tr>
<tr>
<td>TP53</td>
<td>18</td>
<td>Exon 8</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>N/A</td>
<td>Damaging</td>
<td>Damaging</td>
<td>p.Arg263Cys</td>
<td>48</td>
<td>rs149633775</td>
</tr>
<tr>
<td>TP53</td>
<td>5</td>
<td>Exon 8</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly227Trp</td>
<td>64</td>
<td>rs763098116</td>
</tr>
<tr>
<td>TP53</td>
<td>4</td>
<td>Exon 7</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly245Ser</td>
<td>64</td>
<td>rs28304575</td>
</tr>
<tr>
<td>TP53</td>
<td>1</td>
<td>Exon 7</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>p.Arg175His</td>
<td>75</td>
<td>rs28304576</td>
</tr>
<tr>
<td>TP53</td>
<td>8</td>
<td>Exon 5</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>Somatic</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Val157Phe</td>
<td>75</td>
<td>rs121912654</td>
</tr>
<tr>
<td>KRAS</td>
<td>10</td>
<td>Exon 4</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Ala146Thr</td>
<td>65</td>
<td>rs121913527</td>
</tr>
<tr>
<td>KRAS</td>
<td>20</td>
<td>Exon 3</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>p.Thr58Ile</td>
<td>55</td>
<td>rs104894364</td>
</tr>
<tr>
<td>KRAS</td>
<td>3</td>
<td>Exon 2</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>Somatic</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Not applicable</td>
<td>62</td>
<td>rs397517040</td>
</tr>
<tr>
<td>KRAS</td>
<td>5</td>
<td>Exon 2</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly13Asp</td>
<td>64</td>
<td>rs112445441</td>
</tr>
<tr>
<td>KRAS</td>
<td>16</td>
<td>Exon 2</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly13Asp</td>
<td>50</td>
<td>rs112445441</td>
</tr>
<tr>
<td>KRAS</td>
<td>4</td>
<td>Exon 2</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Somatic</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly12Asp</td>
<td>55</td>
<td>rs121913529</td>
</tr>
<tr>
<td>KRAS</td>
<td>11</td>
<td>Exon 2</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly12Mr</td>
<td>48</td>
<td>rs121913529</td>
</tr>
<tr>
<td>KRAS</td>
<td>8</td>
<td>Exon 2</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly12Ser</td>
<td>52</td>
<td>rs121913530</td>
</tr>
<tr>
<td>BRAF</td>
<td>7</td>
<td>Exon 15</td>
<td>SNP</td>
<td>A&gt;A</td>
<td>Het</td>
<td>Somatic/Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Val600Glu</td>
<td>60</td>
<td>rs113488022</td>
</tr>
<tr>
<td>BRAF</td>
<td>19</td>
<td>Exon 15</td>
<td>SNP</td>
<td>C&gt;C</td>
<td>Het</td>
<td>Somatic</td>
<td>Damaging</td>
<td>Damaging</td>
<td>Gly596Val</td>
<td>75</td>
<td>Not provided</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>4</td>
<td>Exon 10</td>
<td>SNP</td>
<td>G&gt;G</td>
<td>Het</td>
<td>Somatic</td>
<td>Damaging</td>
<td>Damaging</td>
<td>p.Glu545Lys</td>
<td>55</td>
<td>rs104886003</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>16</td>
<td>Exon 21</td>
<td>SNP</td>
<td>A&gt;A</td>
<td>Het</td>
<td>Germline</td>
<td>Damaging</td>
<td>Damaging</td>
<td>p.His1047Arg</td>
<td>50</td>
<td>rs121913279</td>
</tr>
</tbody>
</table>

Table 2: Pathogenic and likely pathogenic variants in CRC patients.
Arg147Cys, Arg145His, and Pro116Ser. All of them were in exon 1 and predicted as damaging mutations. Furthermore, PIK3CA showed 5 missense mutations with the following amino acids substitution: Arg524Lys, Gln542Gln, Gln545Lys, Met1043Val, and His1047Arg with one disease-causing nonsense mutation: Arg992Ter (Table 3). BRAF has 2 missense pathogenic variants (Val600Glu and Gly596Val) and another one at the splice site. EGFR has got 2 missense mutations that are predicted as damaging (Gly779Cys patient ID 22 and Val819Met patient ID 19), both of them were found in exon 20. Also, 2 variants were found in splice site in EGFR, one in exon 20 (Ala822Thr) and two in intron 17. AKT1 has got only one variant at splice site: rs770565457, patient ID 17 with no clinical significance. Furthermore, two missense variants were predicted as damaging in exon 5 in GNA11; Arg210Trp in patients IDs 3, 19 and Arg213Trp in patient ID 3. Finally, MET gene has got two missense mutations; a tolerated neutral variant (Gly1102Asp) in exon16 patient ID 3 and a damaging variant (Pro1285Leu) in exon 20 patient ID 21. One variant at splice site was also found; (Asn1209Lys) in exon 18 patient ID 3. In summary, our analysis revealed that most of investigated variants were detected in TP53 (65%) followed by ERBB2 (8.8%), suggesting more involvement of TP53 pathway than KRAS or PI3K pathways in Egyptian population [21].

**TP53 showed the highest percentage of germline inherited mutations among CRC patients**

Total of 37 mutations with known type of inheritance were identified in 24 patients. Out of these 37 mutations, 24 germline mutations were detected (65%), most of them were in TP53 (10/24=42%) followed by KIT (5/24=21%). Thirteen somatic/acquired mutations were identified (35%). Most of the somatic mutations were found equally in KRAS (4/13=30%) and EGFR (4/13=30%). There was no significant correlation between the number of germline mutations in each patient with either lymph node metastasis (P=0.6) or distant metastasis (P=0.78). However, it was noticed that patient ID 4 has got the highest number of both germline mutations (5 mutations); 2 mutations in TP53 and one mutation in each following gene MET, PDGFRA, EGFR in addition to 4 somatic mutations; 2 in PIK3CA, one in KRAS and one in EGFR. Patient ID 4 showed the highest number also of lymph node (LN) involvement (18 LNs) with distant metastasis in liver. No CRC patients who were involved in the current study had positive family history of malignancy except patient ID 14 who was the only one who had family history of uterine, colon and breast cancer. She was 50 years at the diagnosis of CRC with 2 lymph nodes involvement with liver metastasis. Interestingly, this patient did not have a known germline pathogenic variant, but she has got a missense mutation in FOXL2 (G:A, Pro116Ser). This mutation was predicted as damaging by SIFT and Provean, and was not previously described in CRC. Family screening of this variant is recommended for that patient since this missense variant could exist but with much less allelic frequency [31]. Patient ID 14 has got another damaging missense mutation in ERBB2 in exon 18 (G>T, Glu717Asp); this mutation could be an activating mutation that is accompanied with ERBB2 overexpression [32]. Another interesting finding was related to patient ID 9, who was the only case with bone marrow metastasis. Since bone metastasis is extremely rare in CRC [33], we searched for associated variants that might predispose to this sign. No pathogenic variants were detected, besides no characteristic variants were unique to patient ID 9. For instance, G>A, rs1050171 was identified in exon 20; this genotype is a predictor of bad responsiveness to anti-EGFR therapy in metastatic CRC [34]. Also, C>T, rs41736 in exon 20 in MET was also found but with no previous association with bone metastasis (Figure 1). Therefore, cytopathological examination from bone metastasis is very recommended since it is expected to have distinct mutations from that of primary tumour [35]. Concisely, most of mutations detected were germline (65%), commonly present in TP53 (42%) while the somatic mutations (35%) were found mainly in two oncogenes i.e. KRAS and EGFR.

**Patients showed different response to anti-EGFR target therapy**

Finally, we tried to correlate the identified variants with target therapy. Not all patients were subjected to anti-EGFR target therapy since this therapy should be covered by patient’s insurance. Patients IDs 4, 8, 12, 17 and 24 were subjected to avastin. Patients IDs 4, 17 and 24 were good responders with stable disease. Patient ID 4 has got TP53-KRAS- PIK3CA pathogenic variants: TP53 mutation was G>A [rs121912651, Arg248Trp] in exon 7, KRAS pathogenic variant was C>T [rs121913529 Gly12Asp] in exon 20. The latter mutation is one of KRAS mutations that predict anti-EGFR response [36] which appeared in our study to be sensitive to avastin with stability of the disease after applying the drug. The PIK3CA pathogenic variant was G>A rs104886003 Gln545Lys. For the second good responder; patient ID 17 she has got two pathogenic variants in TP53; G>A rs587782529 Arg357Cys in exon 10 and C>T rs28934578 Arg175His in exon 5. However, patient ID 24 did not have any pathogenic variant and the identified variants in this patient were benign, tolerated by SIFT and neutral by Provean predictors. Nevertheless, patient IDs 8 and 12 showed no response and died. For patient ID 8, two pathogenic variants were found; TP53 C>T, rs121912654 Val517Phe in exon 5 and KRAS C>T rs121913530 Gly12Ser in exon 2. However, for patient ID 12 no pathogenic variant was detected only a homozygous TT rs41736 was found in exon 20 in MET gene; This germline mutation was reported in CRC and siblings with 18% allelic frequency [37]. In brief, most of CRC patients who have different pathogenic variants showed variable response to anti-EGFR therapy.

**Discussion**

Basically, CRC is a disease that results from an interplay between...
common and rare variants with different penetrance [38]. This genetic predisposition will assist in screening high-risk family for CRC. Multiple genetic pathways are involved in the pathogenesis of CRC e.g. WNT signalling pathway (adenomatous polyposis coli gene (APC), BubR1), RAS pathway (KRAS, BRAF, mitogen-activated protein kinase MAPK), PI3K pathway with the involved genes: PI3KCA [39]. With the introduction of advanced molecular techniques e.g. next-generation sequencing multiple genetic defects were investigated through genome-wide studies [40-42]. In the present study, we were able to investigate 15 genes generally affected in CRC through TruSight sequencing panel (Illumina). We found 16 pathogenic variants in 12 patients (50% of total patients). Six of the patients participating in this study had TP53 missense pathogenic variants (25%) which are comparable to other population (33%) [43]. However, no one of those six patients met classical Li-Fraumeni syndrome [44] or Li-Fraumeni-like criteria [45]. It was reported that TP53 germline mutations were found in 4-5% in Li-Fraumeni syndrome (LFS). However, presence of TP53 mutations with absence of personal and family history that met LFS criteria was also observed in previous studies [46,47]. Consequently, TP53 had the highest number of investigated mutations including missense, short deletions and splicing mutations as mentioned in results sections; I, a, II and III. Most of them were germline, meaning potentially early onset of the incidence of CRC. Therefore, it is highly recommended to families of those patients to be subjected to direct DNA sequencing of TP53 exons 2-11 as a screening tool [48]. Additionally, eleven patients were found to have KRAS mutations (~46%), which is matching with the published percentage (30% to 50%) [49,50]. The importance of KRAS mutations is its impact on anti-EGFR target therapy in metastatic colon cancer [22] with better overall survival for patients carry wild KRAS mutation than mutant KRAS [51]. Downregulation of KRAS subsequent overexpression of EGFR inhibits PI3K-ERK signalling pathways. Patient ID 4 carried one of the common wild KRAS mutation in codon 12 (G>A, Gly12Asp, G12D, rs121913529) [52]. This patient showed good response to anti EGFR target therapy. That codon 12 KRAS mutation was also detected in patient ID 11. Another common codon 12 KRAS wild mutation (Gly12Ser) was found in patient ID 8 with progression and subsequent death in spite of panitumumab target therapy. Codon 13 wild KRAS common mutation (Gly13Asp) was detected in patient’s ID 5 and 16. Unfortunately, both patients were not subjected to anti-EGFR therapy in order to follow the drug response. Interestingly, patient ID 20 had a wild KRAS mutation outside the hotspot codons (Ala146Thr); this mutation confers phenotype similar to mutations in the hotspot codons thus a similar response to chemotherapy [53]. Another modulator of MAPK pathway is BRAF gene. The coexistence of both KRAS and BRAF is uncommon in CRC patients. In the present study, patients ID 3 and 19 carried 2 different BRAF mutations: Val600Glu (A>T, rs113488022) [54] and Gly596Val (C>A), both variants are pathogenic with different molecular, pathological characteristics and clinical outcomes [55]. These two patients could be benefit from anti-BRAF in combination with anti-EGFR or anti-MAPK [56]. Overall, BRAF mutations (mutant or wild) have worse overall survival than KRAS mutation carriers. Finally, PIK3CA, which plays important role in MAPK pathway, has got pathogenic missense variants in patient ID 4 (G>T, rs104886003, Gln545Lys) and patient ID 16 (A>G, rs121913279, His1047Arg). Both variants are PIK3CA hotspot somatic mutations [57] with good response to cetuximab [25]. Interestingly, patient ID 4 is a very good example of a possible interplay between pathways of TP53, KRAS and PI3K. These patients have 3 pathogenic variants; one in each previous gene and showed a good respond to anti-EGFR therapy. However, patient ID 8, who had also 2 pathogenic variants in PI3 and KRAS,

was bad responder and died. This likely correlation between previous variants could be based on potential molecular cross-talk between different pathways since activation of AKT inhibits TP53-mediated apoptosis. In addition, mutant KRAS antagonizes the downstream of PI3K signalling in colon cancer cell lines [58]. Remarkably, thirteen patients (54% of patients) included in the present study had 11 ERBB2 missense mutations with amino acid substitutions; 8 of them were predicted as damaging disease-causing variants. These variants may be associated with ERBB2 protein overexpression, thus potential response to anti-Her2/neu (trastuzumab) in combination with other agents e.g. cetuximab and pertuzumab [59,60].

Conclusion

To our knowledge, this is the first CRC study performed in Egypt using TruSight colorectal cancer panel. Although the relative small number of patients involved, this study is considered a good start to know the potential variants found in Egypt and the affected genes to be more focused and investigated.

Acknowledgment

Acknowledgment is to Genetic Signature Cancer Center, Tanta University, Egypt

Funding

This work was funded by grants from Tanta University and Ministry of Higher Education, Egypt

References


