UNIVERSITÀ
DEGLI STUDI
DI MILANO

ORIGINAL ARTICLE   **Open Access**

Assessing the genetic impact of *Enterococcus faecalis* infection on gastric cell line MKN74

Check for updates

Seyhan Turk[1], Can Turk[2*], Elif Sena Temirci[3], Umit Yavuz Malkan[4], Gulberk Ucar[1] and Sukru Volkan Ozguven[2]

Abstract

Purpose: *Enterococcus faecalis* (*E. faecalis*) is an important commensal microbiota member of the human gastrointestinal tract. It has been shown in many studies that infection rates with *E. faecalis* in gastric cancer significantly increase. It has been scientifically proven that some infections develop during the progression of cancer, but it is still unclear whether this infection factor is beneficial (reduction in metastasis) or harmful (increase in proliferation, invasion, stem cell-like phenotype) of the host. These opposed data can significantly contribute to the understanding of cancer progress when analyzed in detail.

Methods: The gene expression data were retrieved from Array Express (E-MEXP-3496). Variance, *t* test and linear regression analysis, hierarchical clustering, network, and pathway analysis were performed.

Results: In this study, we identified altered genes involved in *E. faecalis* infection in the gastric cell line MKN74 and the relevant pathways to understand whether the infection slows down cancer progression. Twelve genes corresponding 15 probe sets were downregulated following the live infection of gastric cancer cells with *E. faecalis*. We identified a network between these genes and pathways they belong to. Pathway analysis showed that these genes are mostly associated with cancer cell proliferation.

Conclusion: NDC80, NCAPG, CENPA, KIF23, BUB1B, BUB1, CASC5, KIF2C, CENPF, SPC25, SMC4, and KIF20A genes were found to be associated with gastric cancer pathogenesis. Almost all of these genes are effective in the proliferation of cancer cells, especially during the infection process. Therefore, it is hypothesized that downregulation of these genes may affect gastric cancer pathogenesis by reducing cell proliferation. And, it is predicted that *E. faecalis* infection may be an important factor for gastric cancer.

Keywords: Gastric cancer, *Enterococcus faecalis* infection, *E. faecalis*, Proliferation, Cell cycle Abbreviations, *E. faecalis Enterococcus faecalis*, ROS Reactive oxygen species, CGP Cancer Genome Project, DAVID Database for Annotation, Visualization and Integrated Discovery, GSEA Gene set enrichment analysis, ES Enrichment score, NES Normalized enrichment score, NOM *p* value Nominal *p* value, FDR *q* value False discovery rate *q* value, FWER Familywise error rate *p* value

Introduction

According to the 2018 GLOBOCAN data, gastric cancer has a death rate of 8.2% among ten common cancers and a 5.7% new case rate among five common cancers (Bray et al. 2018). Risk factors include genetics, gastric

ulcer, diet, smoking, and alcohol use, but the main risk factor is bacterial infections (Rawla and Barsouk 2019). Infections are very important in the chemotherapeutic response because of their effects on various mechanisms. Due to weakened immune systems and various treatment procedures, cancer patients are more likely to become infected from the flora.

Most cases of stomach cancer are associated with bacterial infections, particularly *Helicobacter pylori* gastritis,

and have been found to increase the likelihood of stomach cancer by about 6 times within 10 years after infection. However, the data on the roles of bacterial infections in cancer are definitely in two different directions. Some researchers suggest that bacterial infections play a protective role, while others highlight the harmful effects associated with the infection.

Enterococcus is a part of the normal intestinal flora of humans and animals (Dubin and Pamer 2014). *Enterococcus faecalis* (*E. faecalis*) has a fundamental role in the development of intestinal immunity in the early stages of life in the human gastrointestinal system and plays a role as a protective agent in the regulation of colon homeostasis in newborn babies (Fanaro et al. 2003; Are et al. 2008; Wang et al. 2014). Moreover, due to its anti-inflammatory potential, it is considered to be probiotic in the treatment of certain diseases, such as chronic sinusitis, bronchitis, or infant acute diarrhea (Habermann et al. 2001; Huycke et al. 2002; Nueno-Palop and Narbad 2011; Gong et al. 2017). Besides its beneficial effects it is also known to play important roles in the formation of systemic infections and translocations (Khan et al. 2018; Ben Braïek and Smaoui 2019). It can easily colonize in extra-intestinal and unusual areas and can cause infections throughout certain predisposing conditions such as long-term hospitalizations and some immune compromised conditions.

It has been reported in many studies that the increase of *E. faecalis* infections in various cancer types is very common. An increase in *E. faecalis*-related infections has been noted in cancerous lesions of patients with oral and oropharyngeal squamous cell carcinoma (Boonanantanasarn et al. 2012). The increased levels of *E. faecalis* infection and leading to chromosomal imbalance have been shown in colon cancer (Khan et al. 2018). It triggers poor prognosis by causing invasion in bladder cancer cells (Horsley et al. 2013). It has been shown that cells that interact with *E. faecalis* trigger the production of reactive oxygen species (ROS), and the mitochondrial genome is impaired in gastric cancer (Strickertsson et al. 2013).

However, uncontrolled overproduction of ROS can lead to carcinogenic, cytotoxic results. On the contrary, it was found that the colorectal cancer cell line HCT-116-cocultured with *E. faecalis* downregulated the expression of the fasting-induced adipose factor (FIAF) gene, which is normally associated with the development of some cancer types (Grootaert et al. 2011). It has been suggested that *E. faecalis* strains isolated from human breast milk may be candidates for the prevention and treatment of cancer when they inhibit the proliferation of breast cancer cells (Hassan et al. 2016).

These opposed data can contribute significantly to understanding of cancer progression when analyzed in more detail whether the interaction of the *E. faecalis* will

benefit or harm the host. Although there are studies reporting genetic alterations in gastric cancer caused by *E. faecalis* that may affect the chemotherapeutic response and patients' prognosis, withal their roles, related pathways, and mechanisms remains unclear.

In the current study, we aimed to determine the *E. faecalis*-related altered genes, the network relations of these genes with each other and to relevant pathways using all genome profiles of gastric cancer cells. The results can provide important contribution how prognosis can be affected in the presence of infection.

Materials and methods

Microarray gene expression data

The gene expression data was obtained from the Cancer Genome Project (CGP) database (http://www.cancerrxgene.org/). Transcription profile data of human MKN74 gastric epithelial cells treated with *Enterococcus faecalis* were obtained from Array Express (E-MEXP-3496) (Strickertsson et al. 2014).

Processing and normalization of data

The raw data from Array Express were normalized with the Affy package in the R software. Normalized transcription profile data consists of 23,344 different genes/54,675 probe sets. The data contains control day 1, control day 5, *E. faecalis* infected day 1, and *E. faecalis* infected day 5 groups' whole genome expression data, and they were triplicated.

Variance, *t* test, and linear regression analysis

Among the groups, significant genes with a standard deviation value above 1 were identified. In order to group the identified genes more specifically, Pearson's correlation absolute P value based on the correlation coefficient was calculated and genes above 0.05 were selected. In addition, Pearson correlation coefficient value (R value) was calculated and genes above 0.99 were selected. Genes with a P value less than 0.01 were selected by comparing the expression values of the genes with the control groups by two-tailed *t* test analysis.

Analyses were done using GraphPad Prism 5.0 (Graphpad Prism 5 Software, San Diego, CA, USA). To determine the difference between groups two tail *t* test were performed. Genes with a *t* test P value less than 0.01 were selected. P value less than 0.05 were selected in Pearson analysis.

Hierarchical clustering

Genes determined in linear regression analysis were hierarchically clustered with mean standardized gene expression values with the Euclidean Gene Cluster 3.0 program. The data was standardized after cluster analysis, and the standardized data were viewed using Treeview.

Hierarchical clustering was performed by clustering both genes and arrays using Euclidian distance as similarity metric and complete linkage as clustering method.

Network analysis

To generate a network based on coexpression and genetic interactions, the GeneMANIA software were used and genes in similar pathways were identified with Cytoscape (Shannon et al. 2003). Twelve differentially expressed genes from the 138 genes/165 probe sets were used as input. The datasets have been integrated, analyzed, and visualized to find out if they have functionally similar genes associated with each other and to identify related functions for different gene groups in the network. Thus, the network relationship of these genes was determined. The software scores each gene (node). The size of nodes directly correlated with the score. For example, NDC80 gene has virtually 0.56 score and CENPT has 0.05 score. This directly correlated with the size of the nodes. Higher scores indicate genes that are more likely to be functionally related and more likely has coexpression potential.

Pathway enrichment analysis

To understand biological linkage behind these genes, "Database for Annotation, Visualization and Integrated Discovery" (DAVID) software was used. The pathways associated with our genes were identified.

Gene set enrichment analysis (GSEA)

The gene set enrichment analysis (GSEA) was carried out in concordance with GSEA guideline procedure (http://software.broadinstitute.org/gsea/docGSEAUserGuideFrame.html). E-MEXP-3496 data was used in order to perform the analysis. This data includes 54,675 probe sets (23,344 different genes). Analysis was performed between the control day 5 group and the infected day 5 group to understand the pattern of *E. Faecalis* infection among these groups. The main purpose of this analysis is to determine which gene significantly enriched in which gene set that belongs to the GSEA as well as to understand which gene set enriched in which groups.

GSEA calculate the enrichment score (ES), normalized enrichment score (NES), nominal *P* value (NOM *P* value), false discovery rate *q* value (FDR *q* value), and familywise error rate *P* value (FWER). The ES value indicates gene's maximum deviation in gene sets; in other words, this score helps to find most upregulated genes. NES value represents the connection or difference between gene sets and gene expression. The higher NES value shows the elevation of permutations. Hence, higher NES value increase significance of gene sets. In addition to ES and NES values, NOM *P* value evaluates

the importance of ES calculation. Therefore, NOM *P* value directly correlated with ES as well as NES value. Increase of NOM *P* value show critical role of ES. On the other hand, FWER *P* value indicates false positives probability of NES and so, lower FWER *P* value directly and significantly correlated with correctness of NES calculation. Moreover, FDR *q* value is the most vital value of this analysis. This value needs to be lower than 0.25 and even, when this value become smaller, the enrichment of gene sets is more meaningful.

Results

Time-dependent gene expression alterations in the MKN74 gastric cancer cell line treated with *E. faecalis*

Whole genome expression data were analyzed by linear regression to determine gene expression alterations between control day 1, infected day 1, and infected day 5 groups of human gastric carcinoma MKN74 cells. According to the results, 138 genes corresponding to 165 probe sets with a standard deviation value above 1.0, *P* value below 0.05, and Pearson *R* value above 0.99 showed statistically significant expression alteration. For further analyses, we focused on these genes that altered expression diversity between groups. In the time-dependent *E. faecalis*-treated groups, 10 probe sets were positively correlated and upregulated, and 155 probe sets were negatively correlated and downregulated (Supplementary Table 1).

Hierarchical cluster analysis demonstrated gene alterations between 3 certain groups; control day 1, infected day 1, and infected day 5. Results defining as 155 of the probe sets were negatively correlated, highly expressed in the control day 1 group, and expression decreased at infected day 1 group and so on at infected day 5 group. And conversely 10 of the probe sets were positively correlated, low expressed in the control day 1 group, and expression increased at infected day 1 and so on at infected day 5 group. Top 50 genes were visualized in the figure and the rest were given as supplementary data (Fig. 1) (Supplementary Figures 1, 2, 3). So, there were significantly altered genes between control and treated groups.

Gene alterations due to treatment with *E. faecalis*

In order to determine whether this expression change was caused by *E. faecalis* or cancer itself, variance analyses and *t* test were performed between control day 5 groups' and infected day 5 groups' 138 genes/165 probe set expression data. To better analyze gene expression alteration, long range group "infected day 5" group was selected for variance analysis and *t* test. Thus, 12 statistically significant genes (NDC80, NCAPG, CENPA, KIF23, BUB1B, BUB1, CASC5, KIF2C, CENPF, SPC25, SMC4, KIF20A) corresponding 15 probe sets with a standard deviation above 1.0 and a *t* test *P*

Control Day 1
Infected Day 1
Infected Day 5

218662_s_at
211090_s_at
213008_at
1552619_a_at
222457_s_at
215509_s_at
226604_at
204240_s_at
223117_s_at
219493_at
218355_at
228281_at
225834_at
227350_at
203856_at
235088_at
220295_x_at
209464_at
212847_at
201663_s_at
228191_at
226784_at
231855_at
232278_s_at
212037_at
204835_at
220085_at
218396_at
235852_at
219588_s_at
218349_s_at
1555758_a_at
205032_at
202268_s_at
225176_at
232596_at
202833_s_at
209610_s_at
210095_s_at
1555886_at
200799_at
223072_s_at
236400_at
204437_s_at
226403_at
226051_at
1568838_at
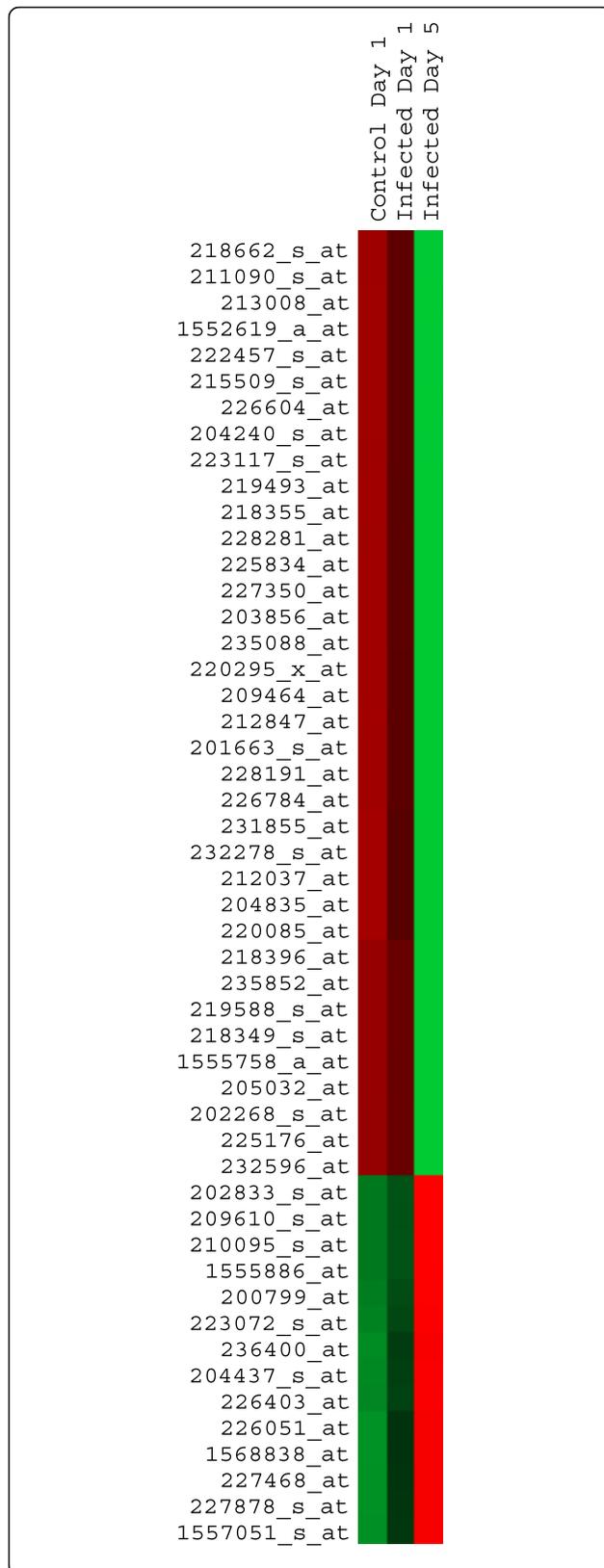227468_at
227878_s_at
1557051_s_at

Fig. 1 Hierarchical clustering of 138 statistically significant genes/165 probe sets in the three groups specified. The analysis reveals sensitive low expressions (green), intermediate (black), and high expressions (red) of 138 genes for control day 1, infected day 1, and infected day 5. Designated groups are clearly classified. Top 50 genes were given in the figure

value less than 0.01 were determined (Table 1) (Supplementary Figure 4).

In linear regression analysis, all of these 12 genes were found to be negatively correlated and downregulated. Figure 2 shows the expression alterations of these 12 genes time-based between control day 1, infected day 1, and infected day 5 (Fig. 2). Gene expressions have been shown to decrease as the infection process progresses in all these 12 genes.

Network construction and identification of key candidate genes using gene-gene interaction

Network analysis was performed with Cytoscape to better demonstrate the biological linkage of these 12 genes whose expression decreased from day 1 to day 5 during infection. As can be seen from the figure, it has been shown that the 12 genes we identified have a strong network relationship between each other and with other candidate genes (Fig. 3). The linking line between genes illuminates the network of these genes. The thickness of the linking line determines the power of connection of the related genes. For instance, the linking line between

Table 1 The list of 12 genes (15 probe sets) which have the most alterations in expression. These genes have standard deviation above 1.0 and *t* test value less than 0.01 between 5 days control group and 5 days treated group. These significant values indicate that the change occurred due to infection.

| Probe Set	Gene Symbol	St Dev	T-Test
204162_at	NDC80	1.76016363	0.00153695
218663_at	NCAPG	1.74829872	0.0002009
204962_s_at	CENPA	1.5109911	0.00039232
218662_s_at	NCAPG	1.46733952	0.00062378
204709_s_at	KIF23	1.42573476	0.00046214
203755_at	BUB1B	1.36025723	1.7849E-05
215509_s_at	BUB1	1.30655845	0.00019534
209642_at	BUB1	1.26836983	0.0001898
228323_at	CASC5	1.19910534	4.6314E-05
211519_s_at	KIF2C	1.17437172	0.00024729
207828_s_at	CENPF	1.13683208	8.8436E-05
209891_at	SPC25	1.13377039	0.00260856
201664_at	SMC4	1.05623047	0.00770427
218755_at	KIF20A	1.02361512	0.0033544
209408_at	KIF2C	1.02269745	0.00046939
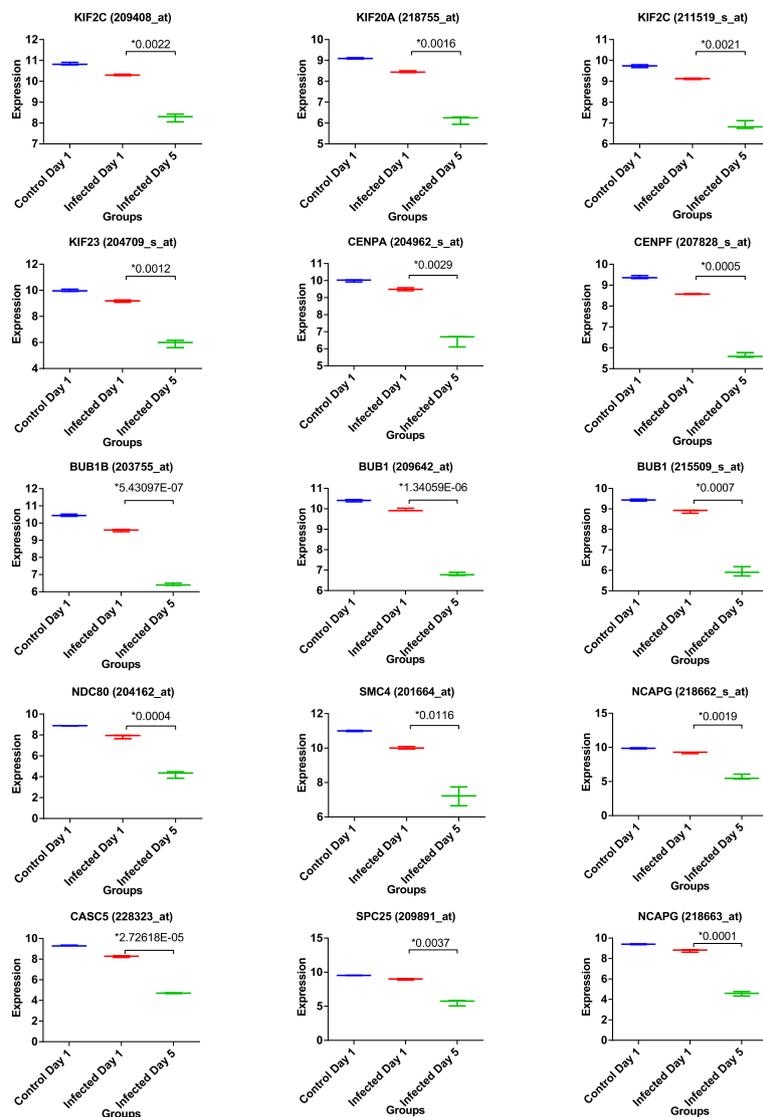
Fig. 2 Twelve genes/15 probe sets that show significant downregulation after treatment with live infection. Statistically significant alterations were detected between control day 1, infected day 1, and infected day 5 for all probe sets
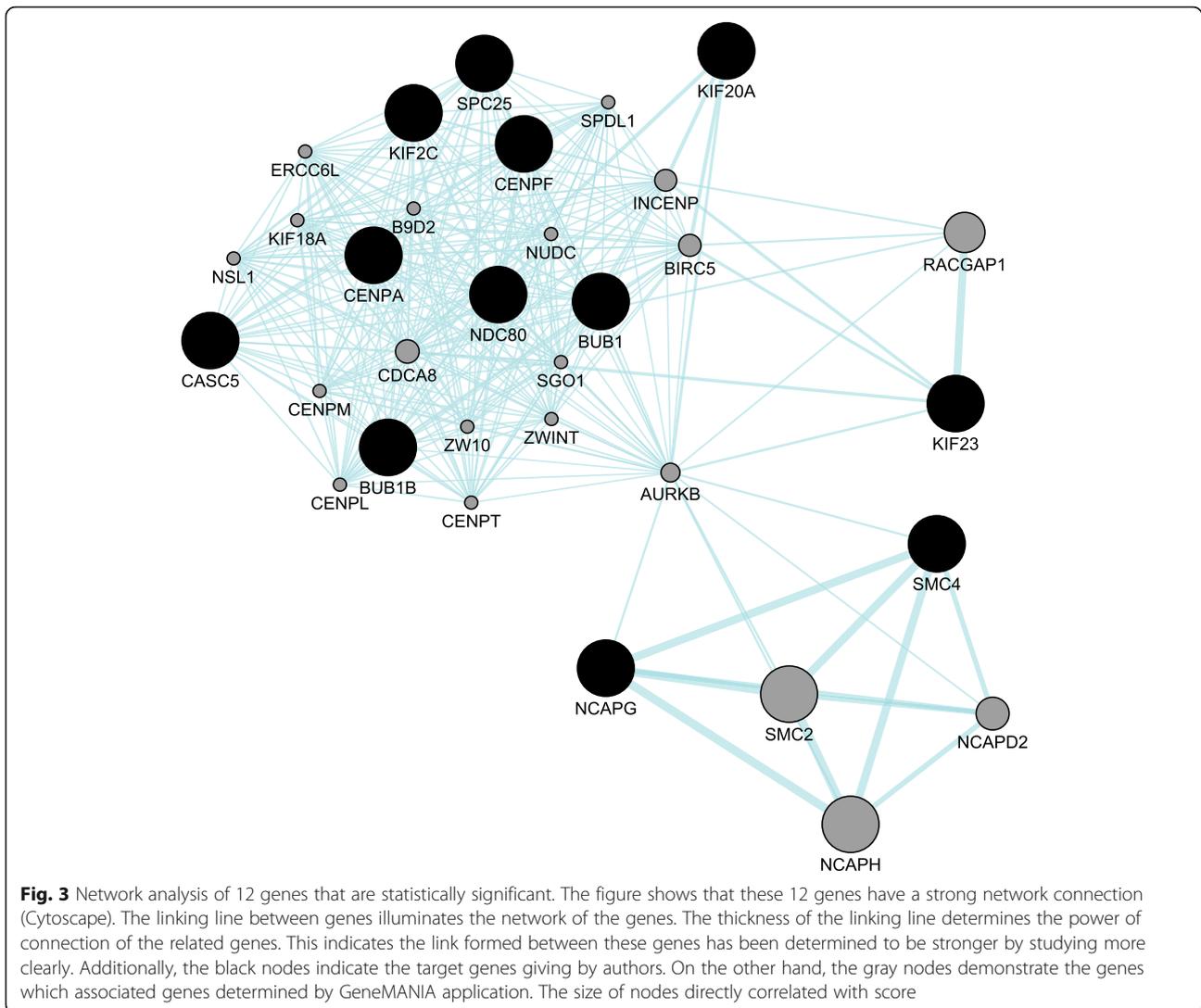
NCAPG and SMC4 represent one of the thickest lines. This indicates the link formed between these genes has been determined to be stronger by studying more clearly. Additionally, the black nodes indicate the target genes giving by authors. On the other hand, the gray nodes demonstrate the genes which associated genes determined by GeneMANIA application.

Functional enrichment of genes and correlations with pathways

Pathway analysis of biological processes was done with DAVID software to reveal the relationship of these 12 genes with cellular functions and pathways and to understand their new meaning. The four important pathways have been identified such as cell cycle, pyrimidine metabolism, mismatch repair, and P53 signaling pathway which are associated with 12 genes during the exposure to *E. faecalis*. The relationship of these pathways with cancer development, such as the importance of cell cycle in gastric cancer cells, is very clear (Table 2).

The significantly enriched gene sets and their ES, NES, NOM *P* value, FWER *P* value, and FDR *q* value represented in Supplementary Table 2. According to GSEA, the cell cycle checkpoint as well as the mitotic cell cycle checkpoint gene set was found as significantly correlated with the genes that are present in the data. These gene sets even correlated with

Turk *et al. Annals of Microbiology*   (2021) 71:8

Page 6 of 10

Fig. 3 Network analysis of 12 genes that are statistically significant. The figure shows that these 12 genes have a strong network connection (Cytoscape). The linking line between genes illuminates the network of the genes. The thickness of the linking line determines the power of connection of the related genes. This indicates the link formed between these genes has been determined to be stronger by studying more clearly. Additionally, the black nodes indicate the target genes giving by authors. On the other hand, the gray nodes demonstrate the genes which associated genes determined by GeneMANIA application. The size of nodes directly correlated with score

pathway analysis in Table 2 and other emphasized studies. The enriched genes were NCD80, BUB1, BUB1B, and CENPF in these two gene sets.

Also, Fig. 4 demonstrated the cell cycle checkpoint gene set plot. This graph showed genes with enriched NCD80, BUB1, BUB1B, and CENPF in the control day 5 group, also called positively correlated. On the other hand, the same genes were downregulated and negatively correlated in the infected day 5 group. The same results were obtained for the mitotic cell cycle checkpoint gene set (Fig. 5).

Discussion

In recent years, the association between intestinal flora-related microbiota and gastric cancer has increased the interest to its roles in human health with crucial findings. *E. faecalis* infections have become important due to their potential to affect cell cycle which can alter the

Table 2 Pathways related to the genes are linked. It is seen that important pathways in cancer progression are related to our genes. Most of these genes are linked to the cell cycle pathway from the database for annotation, visualization, and integrated discovery (DAVID)

Category	Pathways	Count	P value	Benjamini
KEGG_PATHWAY	Cell cycle	6[*1]	1.3E−3	9.0E−2
KEGG_PATHWAY	Pyrimidine metabolism	2[*2]	3.1E−3	1.1E−1
KEGG_PATHWAY	p53 signaling pathway	4[*3]	1.0E−2	1.7E−1
KEGG_PATHWAY	Mismatch repair	3[*4]	9.1E−3	2.0E−1

[*1]NDC80, NCAPG, BUB1B, BUB1, CENPA, KIF23
[*2]BUB1B, BUB1
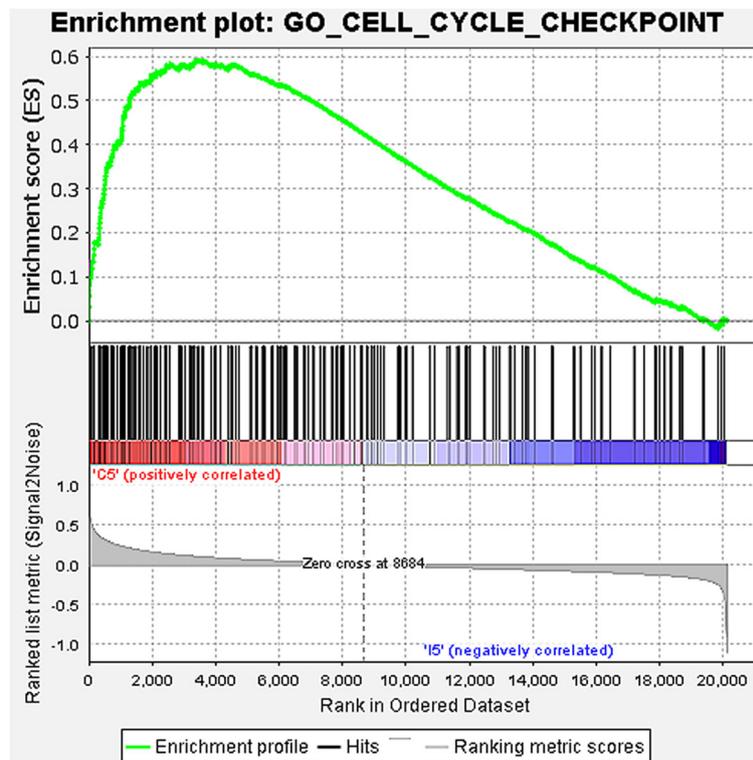[*3]KIF2C, SPC25, KIF20A, CENPF
[*4]SMC4, CENPA, NCAPG

Fig. 4 The cell cycle checkpoint enrichment plot was represented. The black straight line refers the enriched genes in the groups. Red part contains the genes that positively correlated in control day 5 group and were upregulated in control day 5. On contemporarily, the blue line includes downregulated or in other words negatively regulated genes that belong to infected day 5

response to chemotherapy. *E. faecalis* is a part of the normal intestinal flora that can pass through the gut and cause systemic infection (Khan et al. 2018). If the barrier is violated, the intestinal flora such as *E. faecalis* can translocate and cause infection (Ramphal 2004). In this study, it is aimed to identify *E. faecalis-related* gene expression alterations, their network relationships, and associated pathways using all genome expression data of MKN74 gastric cancer cells.

NDC80, NCAPG, CENPA, KIF23, BUB1B, BUB1, CASC5, KIF2C, CENPF, SPC25, SMC4, and KIF20A gene expression alterations were identified depending on time after treatment with *E. faecalis* of gastric cell line MKN70. These genes were identified as downregulated in gastric cancer cells, and the treated and untreated groups ($P < 0.05$) were hierarchically forming a very distinct cluster, as expected.

Gene set enrichment analysis supported the results that some crucial genes function effectively in the cell cycle as a result of *E. Faecalis* infection. NDC80, BUB1B, BUB1, and CENPF genes showed a statistically significant difference in cell cycle checkpoint and mitotic cell cycle checkpoint gene sets associated with cell cycle. Genes found statistically significant as a result of the comparison between the control day 5 and the infected

day 5 were enriched in the specified biological gene sets. This enrichment supports that these genes have important effects on the cycles of cells. In addition, these effects trigger cancer development by causing critical changes in cell function, as indicated.

NDC80 is a kinetocol protein that combines with other proteins to form a microtubule binding site. It has important roles in the cell cycle and an abnormality can lead to apoptotic cell death. Thus, downregulation of this gene may cause abnormalities in the cell cycle of gastric cells after *E. faecalis* infection. In this case, cells can escape from proliferation and go into apoptosis. High expression of the NDC80 gene in pancreatic and breast cancer cells is associated with poor prognosis and plays important roles in tumor formation in pancreatic, breast, and stomach cancer (Meng et al. 2015; Tang and Toda 2015; Liu et al. 2018). NCAPG is the mitotic gene that plays important roles in the cell cycle. High expression of NCAPG is associated with proliferation in gastric cancer (Song et al. 2018). A downregulation of this gene may cause a decrease in proliferation of gastric cancer cells. CENP-A and CENP-F are histone variants that ensure kinetochores and centromeres to form and function properly. High expressions are related with cancer progress and have an important role in cell division (Sun
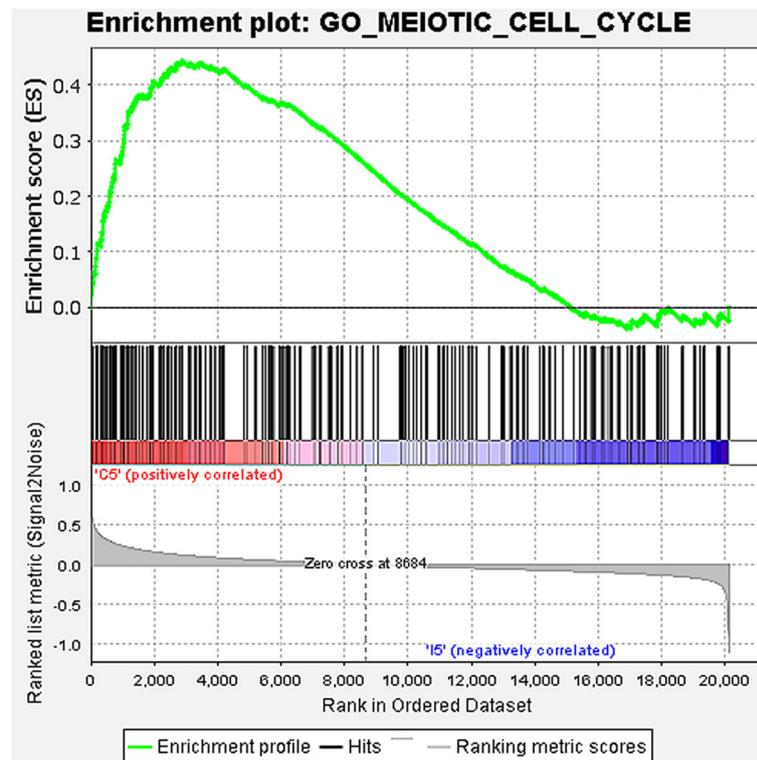
Fig. 5 The mitotic cell cycle checkpoint enrichment plot was represented. The black straight line refers the enriched genes in the groups. Red part contains the genes that positively correlated in control day 5 group and were upregulated in control day 5. On contemporarily, the blue line includes downregulated or in other words negatively regulated genes that belongs to infected day 5

et al. 2016). High expression can be predicted as a biomarker in the malignant progression and poor prognosis of many types of cancer. The decrease in the expression of this gene can cause irregularity in the regulation of the centromere, causing genome imbalance in gastric cancer. Kinesin superfamily proteins (KIFs) are microtubule-dependent motor proteins and function as oncogenes in cancer cells. It was shown that KIF overexpression is associated with poor prognosis and plays important roles in cancer progression and metastasis (Yang et al. 2019). BUB1 is a mitotic spindle checkpoint protein (budding uninhibited by benzimidazoles 1). It has been suggested that high expression of BUB gene family members in gastric cancer correlates strongly with tumor proliferation (Stahl et al. 2017). Its expression was inversely correlated with the residual tumor stage, and low BUB1 expression was associated with shorter survival (Grabsch et al. 2003; Stahl et al. 2017) ($P < 0.001$). CASC5 (cancer susceptibility candidate 5) is required for creation of kinetochore-microtubule attachments and chromosome segregation. It has been shown to play important roles in various types of cancer and has been shown to be a new oncogene in lung adenocarcinoma (Cui et al. 2020). Its downregulation in gastric cancer may be important. Structural maintenance of

chromosome 2 (SMC2) gene encodes condensin complexes and responsible from chromosomal stability (Je et al. 2014). It is showed that SMC2 gene has low expression in gastric cancer, and it is a risk loci for breast and ovarian cancer (Je et al. 2014; Murakami-Tonami et al. 2014; Kar et al. 2016). The SPC25 gene encodes a protein that may be involved in kinetochore-microtubule interaction, and it is shown that upregulation of SPC25 promotes breast cancer (Wang et al. 2019).

It seems that all genes often play a role in the progression of cancer by affecting cell proliferation and especially the cell cycle. An irregularity in the cell cycle pathway is an important factor that could lead gastric tumor formation. Some cyclin-related kinases with bacterial infection are known to upregulate gastric cancer cells. Activation and regular functioning of these cyclin-depended kinases occur due to the progress of the cell cycle (Molaei et al. 2018). Cyclin-depended kinases play an important role in the G1/S phase transition in the cell cycle by targeting E2F transcription factors. An abnormality occurring in this molecular mechanism causes damage to this connection and causes the development of gastric cancer cells.

It has been reported that most of identified genes in this study are highly expressed in various cancers and

Turk *et al. Annals of Microbiology*   (2021) 71:8

Page 9 of 10

are associated with a poor prognosis. Thus, cancer progression may also be expected to decrease due to *E. faecalis* infection. The relationship of some of these genes with gastric cancer was first shown in this study. As a result of network and pathway analysis, it was determined that these genes are associated with important pathways in cancer progression. These pathways are cell cycle, pyrimidine metabolism, mismatch repair, P53 signaling pathway, endocytosis, and spliceosome. It is obvious that most of these pathways are associated with cancer progression and proliferation. And this allows us to positively intervene in the progression of gastric cancer. The limitation of this study is lack of incorporation of in vivo experiments.

Conclusion

In this study, it has been shown that *E. faecalis* infection may have important effects on the proliferation of cells in gastric cancer. The results are important indicators that infections caused by *E. faecalis* directly or indirectly have an effect on the progression of gastric cancer, and this effect may favor the host in fighting with cancer. Most of the genes we identified have functions that slow down cell proliferation. More in vitro and in vivo studies are needed to reveal which structural or functional components are responsible for this adverse effect on gastric cancer cells.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13213-020-01615-3.

Additional file 1: Supplementary Table 1. The list of genes that are significantly expressed as a result of statistical comparison with Pearson analysis. 138 genes (165 Probe sets) are expressed significantly between control day 1, infected day 1 and infected day 5 groups as a result of linear regression analysis. Pearson r correlation analysis with all genom expression data, SD value above 1, r value above 0.99 and p value less than 0.05 were selected.

Additional file 2: Supplementary Table 2. Gene set enrichment analysis (GSEA). GSEA was performed between control day 5 group and infected day 5 group.

Additional file 3: Supplementary Figure 1. Hierarchical clustering of rest of the 138 statistically significant genes / 165 probe sets in the three groups specified. Rest of the genes in the Fig 1. The analysis reveals sensitive low expressions (green), intermediate (black) and high expressions (red) of 138 genes for control day 1, infected day 1 and infected day 5. Designated groups are clearly classified.

Additional file 4: Supplementary Figure 2. Hierarchical clustering of rest of the 138 statistically significant genes / 165 probe sets in the three groups specified. Rest of the genes in the Fig. 1. The analysis reveals sensitive low expressions (green), intermediate (black) and high expressions (red) of 138 genes for control day 1, infected day 1 and infected day 5. Designated groups are clearly classified.

Additional file 5: Supplementary Figure 3. Hierarchical clustering of rest of the 138 statistically significant genes / 165 probe sets in the three groups specified. Rest of the genes in the Fig. 1. The analysis reveals sensitive low expressions (green), intermediate (black) and high

---

expressions (red) of 138 genes for control day 1, infected day 1 and infected day 5. Designated groups are clearly classified.

Additional file 6: Supplementary Figure 4. The 12 genes (15 Probe Sets) which have alterations in expression. These genes have standard deviation above 1.0 and T-test value less than 0.01 between 5 day control group and 5 day infected group. These significant values indicate that the change occurred due to infection. Top 50 genes were visualized in the figure.

Authors' contributions
Sukru Volkan Ozguven and Can Turk conceived and conducted the study. Can Turk, Seyhan Turk, Umit Yavuz Malkan, and Elif Sena Temirci performed the statistical analysis. Can Turk and Seyhan Turk wrote the manuscript. Sukru Volkan Ozguven and Gulberk Ucar revised the manuscript. Sukru Volkan Ozguven and Gulberk Ucar supervised the study. The authors offered intellectual contribution for the manuscript and approved its final form.

Availability of data and materials
A majority of data generated or analyzed during this study included in this published article. The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request. The gene expression data was obtained from the Cancer Genome Project (CGP) database (http://www.cancerrxgene.org/), and transcription profile data were obtained from Array Express (E-MEXP-3496).

Ethics approval and consent to participate
Ethics approval was not necessary because the used in silico data were obtained from Array Express (E-MEXP-3496).

Consent for publication
Not applicable.

Competing interests
The authors declared no competing interests.

Author details
[1]Department of Biochemistry, Faculty of Pharmacy, Hacettepe University, Ankara, Turkey. [2]Department of Medical Microbiology, Faculty of Medicine, Lokman Hekim University, Ankara, Turkey. [3]Department of Molecular Biology and Genetics, Faculty of Science, Bilkent University, Ankara, Turkey. [4]Department of Hematology, Dışkapı Yıldırım Beyazıt Training and Research Hospital, Ankara, Turkey.

References

Are A, Aronsson L, Wang S, Greicius G, Lee YK, Gustafsson JA, Pettersson S, Arulampalam V (2008) Enterococcus faecalis from newborn babies regulate endogenous PPARgamma activity and IL-10 levels in colonic epithelial cells. Proc Natl Acad Sci U S A 105:1943–1948

Ben Braïek O, Smaoui S (2019) Enterococci: between emerging pathogens and potential probiotics. BioMed Res Int 2019:5938210

Boonanantanasarn K, Gill AL, Yap Y, Jayaprakash V, Sullivan MA, Gill SR (2012) Enterococcus faecalis enhances cell proliferation through hydrogen peroxide-mediated epidermal growth factor receptor activation. Infect Immun 80:3545–3558

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68:394–424

Cui Y, Zhang C, Ma S, Guo W, Cao W, Guan F (2020) CASC5 is a potential tumour driving gene in lung adenocarcinoma. Cell Biochem Funct 38(6):733-742.

Dubin K, Pamer EG (2014) Enterococci and their interactions with the intestinal microbiome. Microbiol Spectr 5(6):10.

Fanaro S, Chierici R, Guerrini P, Vigi V (2003) Intestinal microflora in early infancy: composition and development. Acta Paediatr Suppl 91:48–55

Gong J, Bai T, Zhang L, Qian W, Song J, Hou X (2017) Inhibition effect of Bifidobacterium longum, Lactobacillus acidophilus, Streptococcus thermophilus and Enterococcus faecalis and their related products on human colonic smooth muscle in vitro. PLoS One 12:e0189257

Grabsch H, Takeno S, Parsons WJ, Pomjanski N, Boecking A, Gabbert HE, Mueller W (2003) Overexpression of the mitotic checkpoint genes BUB1, BUBR1, and BUB3 in gastric cancer--association with tumour cell proliferation. J Pathol 200:16–22

Grootaert C, Van de Wiele T, Van Roosbroeck I, Possemiers S, Vercoutter-Edouart AS, Verstraete W, Bracke M, Vanhoecke B (2011) Bacterial monocultures, propionate, butyrate and H_2O_2 modulate the expression, secretion and structure of the fasting-induced adipose factor in gut epithelial cell lines. Environ Microbiol 13:1778–1789

Habermann W, Zimmermann K, Skarabis H, Kunze R, Rusch V (2001) The effect of a bacterial immunostimulant (human Enterococcus faecalis bacteria) on the occurrence of relapse in patients with. Arzneimittelforschung 51:931–937

Hassan Z, Mustafa S, Rahim RA, Isa NM (2016) Anti-breast cancer effects of live, heat-killed and cytoplasmic fractions of Enterococcus faecalis and Staphylococcus hominis isolated from human breast milk. In Vitro Cell Dev Biol Anim 52:337–348

Horsley H, Malone-Lee J, Holland D, Tuz M, Hibbert A, Kelsey M, Kupelian A, Rohn JL (2013) Enterococcus faecalis subverts and invades the host urothelium in patients with chronic urinary tract infection. PloS One 8:e83637–e83637

Huycke MM, Abrams V, Moore DR (2002) Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. Carcinogenesis 23:529–536

Je EM, Yoo NJ, Lee SH (2014) Mutational and expressional analysis of SMC2 gene in gastric and colorectal cancers with microsatellite instability. Apmis 122: 499–504

Kar SP, Beesley J, Amin Al Olama A, Michailidou K, Tyrer J, Kote-Jarai Z, Lawrenson K, Lindstrom S, Ramus SJ, Thompson DJ, Kibel AS, Dansonka-Mieszkowska A, Michael A, Dieffenbach AK, Gentry-Maharaj A, Whittemore AS, Wolk A, Monteiro A, Peixoto A, Kierzek A, Cox A, Rudolph A, Gonzalez-Neira A, Wu AH, Lindblom A, Swerdlow A, Ziogas A, Ekici AB, Burwinkel B, Karlan BY, Nordestgaard BG, Blomqvist C, Phelan C, McLean C, Pearce CL, Vachon C, Cybulski C, Slavov C, Stegmaier C, Maier C, Ambrosone CB, Hogdall CK, Teerlink CC, Kang D, Tessier DC, Schaid DJ, Stram DO, Cramer DW, Neal DE, Eccles D, Flesch-Janys D, Edwards DR, Wokozorczyk D, Levine DA, Yannoukakos D, Sawyer EJ, Bandera EV, Poole EM, Goode EL, Khusnutdinova E, Hogdall E, Song F, Bruinsma F, Heitz F, Modugno F, Hamdy FC, Wiklund F, Giles GG, Olsson H, Wildiers H, Ulmer HU, Pandha H, Risch HA, Darabi H, Salvesen HB, Nevanlinna H, Gronberg H, Brenner H, Brauch H, Anton-Culver H, Song H, Lim HY, McNeish I, Campbell I, Vergote I, Gronwald J, Lubinski J, Stanford JL, Benitez J, Doherty JA, Permuth JB, Chang-Claude J, Donovan JL, Dennis J, Schildkraut JM, Schleutker J, Hopper JL, Kupryjanczyk J, Park JY, Figueroa J, Clements JA, Knight JA, Peto J, Cunningham JM, Pow-Sang J, Batra J, Czene K, Lu KH, Herkommer K, Khaw KT, Matsuo K, Muir K, Offitt K, Chen K, Moysich KB, Aittomaki K, Odunsi K, Kiemeney LA, Massuger LF, Fitzgerald LM, Cook LS, Cannon-Albright L, Hooning MJ, Pike MC, Bolla MK, Luedeke M, Teixeira MR, Goodman MT, Schmidt MK, Riggan M, Aly M, Rossing MA, Beckmann MW, Moisse M, Sanderson M, Southey MC, Jones M, Lush M, Hildebrandt MA, Hou MF, Schoemaker MJ, Garcia-Closas M, Bogdanova N, Rahman N, Le ND, Orr N, Wentzensen N, Pashayan N, Peterlongo P, Guenel P, Brennan P, Paulo P, Webb PM, Broberg P, Fasching PA, Devilee P, Wang Q, Cai Q, Li Q, Kaneva R, Butzow R, Kopperud RK, Schmutzler RK, Stephenson RA, MacInnis RJ, Hoover RN, Winqvist R, Ness R, Milne RL, Travis RC, Benlloch S, Olson SH, McDonnell SK, Tworoger SS, Maia S, Berndt S, Lee SC, Teo SH, Thibodeau SN, Bojesen SE, Gapstur SM, Kjaer SK, Pejovic T, Tammela TL, Dork T, Bruning T, Wahlfors T, Key TJ, Edwards TL, Menon U, Hamann U, Mitev V, Kosma VM, Setiawan VW, Kristensen V, Arndt V, Vogel W, Zheng W, Sieh W, Blot WJ, Kluzniak W, Shu XO, Gao YT, Schumacher F, Freedman ML, Berchuck A, Dunning AM, Simard J, Haiman CA, Spurdle A, Sellers TA, Hunter DJ, Henderson BE, Kraft P, Chanock SJ, Couch FJ, Hall P, Gayther SA, Easton DF, Chenevix-Trench G, Eeles R, Pharoah PD, Lambrechts D (2016) Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types. Cancer Discov 6:1052–1067

Khan Z, Siddiqui N, Saif MW (2018) Enterococcus faecalis infective endocarditis and colorectal carcinoma: case of new association gaining ground. Gastroenterol Res 11:238–240

Liu X, Wu J, Zhang D, Bing Z, Tian J, Ni M, Zhang X, Meng Z, Liu S (2018) Identification of potential key genes associated with the pathogenesis and prognosis of gastric cancer based on integrated bioinformatics analysis. Front Genet 9:265–265

Meng Q-C, Wang H-C, Song Z-L, Shan Z-Z, Yuan Z, Zheng Q, Huang X-Y (2015) Overexpression of NDC80 is correlated with prognosis of pancreatic cancer and regulates cell proliferation. Am J Cancer Res 5:1730–1740

Molaei F, Forghanifard MM, Fahim Y, Abbaszadegan MR (2018) Molecular signaling in tumorigenesis of gastric cancer. Iran Biomed J 22:217–230

Murakami-Tonami Y, Kishida S, Takeuchi I, Katou Y, Maris JM, Ichikawa H, Kondo Y, Sekido Y, Shirahige K, Murakami H, Kadomatsu K (2014) Inactivation of SMC2 shows a synergistic lethal response in MYCN-amplified neuroblastoma cells. Cell Cycle (Georgetown, Tex) 13:1115–1131

Nueno-Palop C, Narbad A (2011) Probiotic assessment of Enterococcus faecalis CP58 isolated from human gut. Int J Food Microbiol 145:390–394

Ramphal R (2004) Changes in the etiology of bacteremia in febrile neutropenic patients and the susceptibilities of the currently isolated pathogens. Clin Infect Dis 39(Suppl 1):S25–S31

Rawla P, Barsouk A (2019) Epidemiology of gastric cancer: global trends, risk factors and prevention. Prz Gastroenterol 14:26–38

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504

Song B, Du J, Song DF, Ren JC, Feng Y (2018) Dysregulation of NCAPG, KNL1, miR-148a-3p, miR-193b-3p, and miR-1179 may contribute to the progression of gastric cancer. Biol Res 51:44

Stahl D, Braun M, Gentles AJ, Lingohr P, Walter A, Kristiansen G, Gütgemann I (2017) Low BUB1 expression is an adverse prognostic marker in gastric adenocarcinoma. Oncotarget 8:76329–76339

Strickertsson JA, Desler C, Martin-Bertelsen T, Machado AM, Wadstrom T, Winther O, Rasmussen LJ, Friis-Hansen L (2013) Enterococcus faecalis infection causes inflammation, intracellular oxphos-independent ROS production, and DNA damage in human gastric cancer cells. PLoS One 8:e63147

Strickertsson JA, Rasmussen LJ, Friis-Hansen L (2014) Enterococcus faecalis infection and reactive oxygen species down-regulates the miR-17-92 cluster in gastric adenocarcinoma cell culture. Genes (Basel) 5:726–738

Sun X, Clermont PL, Jiao W, Helgason CD, Gout PW, Wang Y, Qu S (2016) Elevated expression of the centromere protein-A(CENP-A)-encoding gene as a prognostic and predictive biomarker in human cancers. Int J Cancer 139:899–907

Tang NH, Toda T (2015) MAPping the Ndc80 loop in cancer: a possible link between Ndc80/Hec1 overproduction and cancer formation. Bioessays 37:248–256

Wang Q, Zhu Y, Li Z, Bu Q, Sun T, Wang H, Sun H, Cao X (2019) Up-regulation of SPC25 promotes breast cancer. Aging 11:5689–5704

Wang S, Hibberd ML, Pettersson S, Lee YK (2014) Enterococcus faecalis from healthy infants modulates inflammation through MAPK signaling pathways. PLoS One 9:e97523

Yang Z, Li C, Yan C, Li J, Yan Y, Liu B, Zhu Z, Wu Y, Gu Q (2019) KIF14 promotes tumor progression and metastasis and is an independent predictor of poor prognosis in human gastric cancer. Biochim Biophys Acta Mol Basis Dis 1865:181–192

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.