Development of a Custom NGS Panel for the Determination of Bladder Cancer Risk

Imen Hemissi
INSAT: Institut National des Sciences Appliquees et de Technologie

SAMI BOUSETTA
Universite de Tunis El Manar Faculte des Sciences de Tunis

Hamza Dallali
IPT: Institut Pasteur de Tunis

Faycel Hellal
INSAT: Institut National des Sciences Appliquees et de Technologie

Geoffroy Durand
CIRC: International Agency for Research on Cancer

Catherine Voegele
CIRC: International Agency for Research on Cancer

Haroun AYED
Charles Nicolle Hospital: Hopital Charles Nicolle

Selim Zaghbib
Charles Nicolle Hospital: Hopital Charles Nicolle

Zeineb Naimi
Institut Salah-Azaïz: Institut Salah-Azaiz

Mouna Ayadi
Institut Salah-Azaïz: Institut Salah-Azaiz

Mohamed Chebil
Charles Nicolle Hospital: Hopital Charles Nicolle

James Mckay
IARC: International Agency for Research on Cancer

Florence Le Calvez-Klem
IARC: International Agency for Research on Cancer

Slah Ouerhani (slah_mekni@yahoo.fr)
laboratoire ingénierie des protéines et des molécules recombinantes. Institut National des Sciences Appliquees et de Technologie de Tunis
https://orcid.org/0000-0002-3651-4690

Research Article

Keywords: Bladder cancer, NGS, Panel, SNP, tobacco, Tunisia

DOI: https://doi.org/10.21203/rs.3.rs-840773/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Bladder cancer (BCa) is a heterogeneous disease caused by the interaction between environmental and genetic risk factors. The objective of this study was to design a panel that evaluates the role of some selected variants in BCa susceptibility. We are also interested in studying the interaction between environmental and genetic risk factors.

Methods

The case-controls cohort was composed with 249 BCa cases and 255 controls. The designed Bladder cancer hereditary panel (BCHP) is composed of 139 selected variants. These variants were genotyped by an amplification-based targeted Next-Generation Sequencing (NGS) on the Ion Torrent Proton sequencer (Life Technologies, Ion Torrent technology).

Results

We have found that rs162555, rs2228000, rs10936599, rs710521, rs3752645, rs804276, rs4639, rs4881400 and rs288980 were significantly associated with decreased risk of bladder cancer. However the homozygous genotypes for VPS37C (rs7104333, A/A), MPG (rs1013358, C/C) genes or the heterozygous genotype for ARNT gene (rs1889740, rs2228099, rs2256355, rs2864873), GSTA4 (rs17614751) and APOBR/IL27 (rs17855750) were significantly associated with increased risk of bladder cancer development compared to reference group (OR=2.53, 2.34, 1.99, 2.00, 2.00, 1.47, 1.96 and 2.27 respectively). We have also found that non-smokers patients harboring heterozygous genotypes for ARNT/ rs2864873 (A>G), ARNT/ rs1889740 (C>T) or GSTA4/rs17614751 (G to A) were respectively at 2.775, 3.069 and 6.608 –folds increased risk of BCa development compared to non-smokers controls with wild genotypes. Moreover the ARNT CT (rs1889740), ARNT CG (rs2228099), ARNT TC (rs2864873) and GSTA4 genotypes were associated with an increased risk of BCa even in absence of professional risk factors. Finally the decision-tree analysis produced a three major BCa class. These three classes were essentially characterized by an intensity of tobacco use more than 20 pack years (PY) and the CYP1A2 (rs762551) genotype.

Conclusions

The determined association between genetic variations in BCa and environmental factors, as well as the effect of studied pathway SNPs in comparison with environmental exposition may provide urologists additional genetic information that may help for clinical assessment and treatment decisions. Nevertheless, the underlying mechanisms through which these genes or SNPs affect the clinical behavior of BCAs require further studies.

Highlight

In this study we determined the effect of studied pathway SNPs in comparison with environmental exposition using Next-Generation Sequencing (NGS). We have found that homozygous minor genotype or heterozygous genotype for rs1889740 C/T, rs2228099 C/G, rs2256355 T/C, rs2864873 A/G in ARNT gene (Aryl hydrocarbon receptor nuclear translocator) were significantly associated to increased risk of bladder cancer even in the absence of tobacco risk factors, or professional risk factors. Finally we noticed the importance of tobacco status and the CYP1A2 (rs762551) genotype in the stratification of Tunisian BCa patients.

Background

Bladder cancer (BCa) is a common malignancy of the urinary tract and the ninth most frequent cancer in the world, affecting nearly 3.4 million people with 430 000 new cases diagnosed in 2012 and 165 000 deaths per year (2% of all the cases)[1]. In the few last years, the incidence of BCa was higher in developed countries and some African countries, with an increase in the aggressiveness of the disease and the mortality rate [2]. According to the Cancer Registry of the north of Tunisia (1999–2003), BCa is the most frequent urological cancer in Tunisia in men and in 2018 it became the second malignant tumor with an age-standardized incidence rates of 17.7 and 2.0 per 100 000 in men in women respectively[3]. In 2018, the number of new BCa cases was 1323 [4]. The mean age at diagnosis is 65.9 years, without sex significant difference. At the time of diagnosis, non-muscle-invasive BCa (NMIBC) represents the majority of BCa whereas muscle-invasive BCa (MIBC) represents around 25% of bladder tumors with a 5-year survival rate ranging from 35–70%. NMIBC is usually associated with a favorable prognosis but it is characterized by a variable rate of recurrence and progression depending on tumor characteristics. MIBC is, however, initially aggressive and given the associated high risk of developing a metastatic disease requires a radical treatment when still localized[5].

BCa is a heterogeneous disease associated with many risk factors. The most known risk factor is tobacco smoking, which accounts for the occurrence of up 50% of BCa [6]. Occupational exposure to chemicals, genetic factors, and other environmental factors such as dietary factors, lifestyle, medical condition, fluid intake, also contribute to BCa carcinogenesis, although the evidence of the role of some of these risk factors is still inconclusive [6, 7]. Most carcinogens undergo activation by Phase I enzymes, mostly through oxidation reactions, and detoxification by Phase II enzymes, which is the case for aromatic and heterocyclic amines, well-known BCa carcinogens present in smoking and occupational exposure [8]. Therefore, early studies evaluated whether candidate genetic variants in such metabolic pathways could modify the risk of BCa induced by the above-mentioned carcinogens. For example, the deletion of xenobiotic metabolism gene “GSTM1 and/or GSTT1” characterized as GSTM1- and GSTT1-null genotypes have been shown to be BCa susceptibility loci in several studies [9, 10]. Moreover, N-acetyltransferase 2 (NAT2) and Glutathione-s-transferase M1 (GSTM1) variants were the first described genes showing a gene-smoking interaction associated with increased risk of BCa[11]. Hypothetically the deficiency in xenobiotic metabolizing pathway could result in the accumulation of somatic mutations which should be corrected by DNA-repair pathways. There are five major important DNA-repair pathways consisting of
more than 130 genes: nucleotide excision repair (NER), base excision repair (BER), mismatch repair (MMR), double-strand break repair (DSBR) and transcription-coupled repair (TCR) \([12–14]\). Among these pathways, NER is the most important DNA repair mechanism responsible for various types of DNA damage consisting of oxidative DNA damage, bulky adducts cross-links, alkylating damage, and thymidine dimmers\([15]\). Deficient DNA repair capacity is known to be a cancer-predisposing factor. Indeed the presence of some single nucleotide polymorphisms (SNP) in DNA repair genes can impair the function of the repair enzymes, thus reducing DNA Repair Capacity (DRC) and inducing genetic instability \([16]\). In addition others studies have reported a significant association between variations in gene encoded for cell cycle and/or inflammatory response proteins and bladder development and/or prognosis\([17]\).

All of this reported information highlights the complexity of the mechanisms implicated in the etiology of BCa which warrants further investigation. The goal of this case-control study was to evaluate the implication a selected SNP panel in BCa development of the Tunisian population. We developed an amplification-based targeted Next-Generation Sequencing (NGS) assay to simultaneously genotype 139 polymorphisms in 97 genes, selected according to their eventual implications in the etiology and the prognosis BCa such as those located in genes encoded for enzymes implicated in xenobiotic metabolism, DNA repair, treatment response, and inflammatory reaction process. The design of the panel was done, after consulting published data and available databases (dbSNP, UCSC Genome Browser, and GWAS).

**Methods**

### Study population

A case-control study was conducted in the Urology department of Charles Nicolle Hospital in Tunis (Tunisia) and the Laboratory of Protein Engineering and Bioactive Molecules (LIP-MB 11ES24) of the National Institute of Applied Sciences and Technology of Tunis (INSAT), in collaboration with the Genetic Cancer Susceptibility group (GCS group) of the International Agency for Research on Cancer in Lyon (IARC/WHO). Recruitment was carried out after the agreement of the ethics committee of the Charles Nicolle Hospital and approved by an Ethics Committee (IEC Project No. 17–35) and an MTA (MATERIAL TRANSFER AGREEMENT MTA/ 2017 / IMP / GCS/ 0356) from IARC.

This study was performed on a cohort of 504 cases (249 BCa patients and 255 controls). A total of 255 control cases, randomly selected from healthy volunteer cases enrolled in the Biochemistry Department, were matched to those of the BCa patients group according to the age range (64.28 ± 11.54), sex, and geographic origin. Epidemiological, clinical and anatomopathological data of BCa patients were collected from the medical records of the urology department and the histological reports from the pathology department of the Charles Nicolle Hospital in Tunis (Table 1). Epidemiological data (age, sex, smoking status) and pathological data (grade, stage, histological type, type, and follow-up of treatment) were recorded and made available for the study (Table 1). Histological reviews of the tumors were performed confirmed by a trained pathologist in urological oncology. The BCa tumors of the 249 patients were classified as 45 Specimens of High-Grade (HG) NMIBC, 119 specimens of Low-Grade (LG) NMIBC, and 66 specimens of MIBC.

**Table 1:** Clinical and Epidemiological characteristics of bladder cancer patients
Clinical and epidemiological parameters

|                                | Bladder cancer patients | Controls |
|--------------------------------|--------------------------|----------|
| **Samples sizes**              |                          |          |
| Male                           | 249                      | 255      |
| Female                         | 224 (89.96%)             | 196 (76.86%) |
| Female                         | 25 (10.04%)              | 59 (23.14%) |
| **Mean Age at diagnosis (years)** | 68.18 ± 12.80            | 64.28 ± 11.54 |
| **Smoking status**             |                          |          |
| Smokers                        | 199 (79.91%)             | 138 (54.12%) |
| Non smokers                    | 27 (10.84%)              | 117 (45.88%) |
| ND                             | 23 (9.25%)               | 0 (00.00%) |
| **Number of pack/years**       |                          |          |
| < 20 PY                        | 31 (15.58%)              | 70 (50.72%) |
| ≥ 20 PY                        | 168 (84.42%)             | 68 (49.28%) |
| **Exposure to professional risk factors (farmer, painter, building, chemical factory...)** | | |
| Not exposed                    | 119 (47.80%)             | 225 (88.24%) |
| Exposed                        | 83 (33.33%)              | 30 (11.76%) |
| ND                             | 47 (18.87%)              | 0 (00.00%) |
| **TNM classification**         |                          |          |
| LG NMIBC                       | 119 (47.80%)             | -        |
| HG NMIBC                       | 45 (18.07%)              | -        |
| MIBC                           | 66 (27.50%)              | -        |
| ND                             | 19 (07.63%)              | -        |
| ND: Not Determined; PY: Packet per Year; LG: Low Grade; HG: Haut Grade; MIBC: Muscle Invasive Bladder Cancer; NMIBC: Non-Muscle Invasive Bladder Cancer

Biological Samples And DNA Extraction

Peripheral blood samples were collected into tubes with ethylene diamine tetra-acetic acid EDTA (PH 8). Genomic DNA was extracted from leukocytes using a phenol / chloroform procedure. First, the integrity of the genomic DNA was visualized by electrophoresis on a 1% agarose gel stained with ethidium bromide. In a second step, DNA samples were quantified by Nanodrop and Qubit (High Sensitivity HS / Broad Rang BR) and were dried with the speedvac (Eppendorf” Vacufuge™ Concentrator) to be normalized into 6*96-well plates.

Panel Design

We have selected a panel of 139 polymorphisms from 97 genes. These SNP were selected according to their eventual implications in poor-prognosis BCa such as those affecting the xenobiotic metabolism, DNA repair, treatment response, and inflammatory reaction. The design of the panel was done, after consulting published data and available databases (dbSNP, UCSC Genome Browser, and GWAS).

In a second step, the implication of the SNPs selected in our SNP panel in BCa susceptibility was confirmed by the following tools: Kyoto Encyclopedia of Genes and Genomes (KEGG) available on the website https://www.genome.jp/kegg/pathway.html [18, 19], Panther classification system available on the website (http://pantherdb.org/tools/gxIdsList.do?list=upload_1&organism=Homo%20sapiens)[20], STRING database available on the website https://string-db.org/AmiGO2[21], Reactome Pathway Database[22] and GWAS [23]. In addition, we constructed candidate pathways based on our defined panel consisting of 97 genes. The pathways were defined according to the STRING protein–protein interaction networks. A complete list of the studied variants was summarized in Table 2. Pathways and protein-protein interaction, in which the studied genes were implicated, were presented and described in Fig. 2.
| CHR | SNP      | GENES          | VARIATION                          | Reference variant | Mutated variant | Major allele/Minor allele | Major.allele.freq | Allelic P  |
|-----|----------|----------------|------------------------------------|-------------------|----------------|--------------------------|------------------|-----------|
| 1   | rs2020902| CASP9          | splice_region_variant, intron_variant | A                 | G              | A/G                      | 82,5             | 0,873     |
| 1   | rs2647396| BCL10          | intron_variant                      | C                 | T              | T/C                      | 65,3             | 0,685     |
| 1   | rs560018  | GSTM4          | intron variant                      | T                 | C              | T/C                      | 82,8             | 0,938     |
| 1   | rs11101992| GSTM5          | intron_variant, non_coding_transcript_variant | A                 | C              | A/C                      | 65,9             | 0,596     |
| 1   | rs4970774| GSTM5          | intron_variant, non_coding_transcript_variant | A                 | C              | C/A                      | 50,1             | 0,754     |
| 1   | rs4970776| GSTM5          | intron_variant, non_coding_transcript_variant | T                 | A              | T/A                      | 53,3             | 0,87      |
| 1   | rs1571858| GSTM3          | intron_variant                      | C                 | T              | C/T                      | 75,9             | 0,888     |
| 1   | rs15864  | EPS8L3         | 3_prime_UTR_variant                 | G                 | C              | G/C                      | 81,6             | 0,947     |
| 1   | rs3136701| CD2            | downstream_gene_variant             | G                 | C              | C/G                      | 52,8             | 0,79      |
| 1   | rs1889740| ARNT           | intron_variant                      | C                 | T              | C/T                      | 57,8             | 0,207     |
| 1   | rs2228099| ARNT           | synonymous_variant                  | C                 | G              | C/G                      | 57,4             | 0,203     |
| 1   | rs1027699| ARNT           | intron_variant                      | T                 | C              | T/C                      | 64,8             | 0,937     |
| 1   | rs2256355| ARNT           | intron_variant                      | T                 | C              | T/C                      | 56,7             | 0,181     |
| 1   | rs2864873| ARNT           | intron_variant                      | A                 | G              | A/G                      | 64,1             | 0,78      |
| 1   | rs763110 | FASLG          | upstream_gene_variant               | C                 | T              | T/C                      | 58,6             | 0,998     |
| 1   | rs228001 | ASTN1          | intron_variant, non_coding_transcript_variant | C                 | A              | C/A                      | 96               | 0,937     |
| 1   | rs1800890| IL10           | upstream_gene_variant               | A                 | T              | A/T                      | 74,9             | 0,886     |
| 1   | rs285461 | EPHX1          | intron_variant                      | C                 | A              | C/A                      | 73,8             | 0,223     |
| 1   | rs1051740| EPHX1          | missense_variant                    | T                 | C              | T/C                      | 72,5             | 0,66      |
| 1   | rs2260863| EPHX1          | intron_variant                      | G                 | C              | C/G                      | 64,9             | 0,89      |
| 1   | rs1805410| PARP1          | intron_variant                      | T                 | C              | T/C                      | 82,7             | 0,432     |
| 2   | rs391835 | AC012593.1/CCR2| intron_variant, non_coding_transcript_variant | G                 | A              | A/G                      | 53,5             | 0,93      |
| 2   | rs1056836| CYP1B1/RMDN2   | downstream_gene_variant             | G                 | C              | C/G                      | 56,1             | 0,618     |
| 2   | rs162555 | CYP1B1         | upstream_gene_variant               | T                 | C              | T/C                      | 76,1             | 0,017     |
| 2   | rs3771171| IL18R1         | intron_variant                      | T                 | C              | T/C                      | 80               | 0,165     |

Chr: Chromosome; SNP: Single Nucleotide Polymorphism; P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval
| CHR | SNP       | GENES   | VARIATION           | Reference variant | Mutated variant | Major allele / Minor allele | Major allele freq | Allelic # |
|-----|-----------|---------|---------------------|------------------|----------------|-----------------------------|-----------------|----------|
| 2   | rs11892031| UGT1A10 | intron_variant      | A                | C              | A/C                         | 90.6            | 0.213    |
| 3   | rs2228000 | XPC     | missense_variant    | G                | A              | G/A                         | 81.5            | 0.024    |
| 3   | rs1050450 | RHOA    | downstream_gene_variant | G          | A              | G/A                         | 74.6            | 0.953    |
| 3   | rs10936599| ACTRT3  | upstream_gene_variant | C            | T              | C/T                         | 85.2            | 0.004    |
| 3   | rs710459  | MASP1   | downstream_gene_variant | G            | A              | G/A                         | 71.6            | 0.968    |
| 3   | rs710521  | TP66    | regulatory_region_variant | T            | C              | T/C                         | 76.7            | 0.035    |
| 4   | rs798766  | TACC3   | intron_variant      | T                | C              | C/T                         | 69.7            | 0.836    |
| 4   | rs4073    | IL8     | upstream_gene_variant | A                | T              | A/T                         | 55.3            | 0.83     |
| 4   | rs2227306 | IL8     | intron_variant      | C                | T              | C/T                         | 68.5            | 0.428    |
| 4   | rs3804099 | TLR2    | synonymous_variant  | T                | C              | T/C                         | 53              | 0.259    |
| 4   | rs3087455 | CASP3   | intron_variant      | T                | G              | T/G                         | 65              | 0.4      |
| 5   | rs34847072| AHRR    | 3_prime_UTR_variant | G                | C              | G/C                         | 82.9            | 0.106    |
| 5   | rs401681  | TERT/CLPTM1L | intron_variant     | C                | T              | T/C                         | 53.8            | 0.415    |
| 5   | rs2070874 | IL4     | 5_prime_UTR_variant | C                | T              | C/T                         | 80.9            | 0.455    |
| 6   | rs4510656 | CDKAL1  | intron_variant      | C                | A              | C/A                         | 61.9            | 0.103    |
| 6   | rs451774  | GPX5    | 3_prime_UTR_variant | A                | G              | A/G                         | 60.5            | 0.377    |
| 6   | rs8193036 | IL17A   | upstream_gene_variant | C            | T              | T/C                         | 74.3            | 0.9      |
| 6   | rs2275913 | IL17A   | upstream_gene_variant | G            | A              | G/A                         | 79.8            | 0.207    |
| 6   | rs2180314 | GSTA2   | missense_variant    | C                | G              | G/C                         | 54.9            | 0.398    |
| 6   | rs2144698 | GSTA2   | intron_variant      | T                | G              | G/T                         | 79.5            | 0.679    |
| 6   | rs17614751| GSTA4   | downstream_gene_variant | G          | A              | G/A                         | 95.4            | 0.029    |
| 6   | rs367836  | GSTA4   | downstream_gene_variant | G            | T              | T/G                         | 59.7            | 0.764    |
| 6   | rs150126  | MAP3K7  | intron_variant      | C                | T              | T/C                         | 74.9            | 0.999    |
| 6   | rs4880    | SOD2    | missense_variant    | A                | G              | A/G                         | 53              | 0.406    |

Chr: Chromosome; SNP: Single Nucleotide Polymorphism; P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval
| CHR | SNP     | GENES     | VARIATION                  | Reference variant | Mutated variant | Major allele/Minor allele | Major.allele.freq | Allelic # |
|-----|---------|-----------|-----------------------------|-------------------|-----------------|--------------------------|-------------------|----------|
| 7   | rs1062492 | SNX8      | downstream_gene_variant     | C                 | T               | C/T                      | 79.8              | 0.234    |
| 7   | rs2017000 | EGFR      | intron_variant              | A                 | G               | A/G                      | 74.5              | 0.139    |
| 7   | rs1140475 | EGFR      | synonymous_variant          | T                 | C               | C/T                      | 92.3              | 0.59     |
| 7   | rs2293347 | EGFR      | synonymous_variant          | C                 | T               | C/T                      | 96                | 0.698    |
| 7   | rs1045642 | ABCB1     | synonymous_variant          | A                 | G               | G/A                      | 69.1              | 0.12     |
| 7   | rs1858923 | ABCB1     | intron_variant              | A                 | G               | A/G                      | 61.8              | 0.611    |
| 7   | rs2740574 | CYP3A4    | upstream_gene_variant       | C                 | T               | T/C                      | 79.4              | 0.129    |
| 7   | rs1773597 | CYP3A4    | upstream_gene_variant       | G                 | C               | G/C                      | 92.3              | 0.247    |
| 7   | rs3752645 | PRKAR2B   | intron_variant              | G                 | A               | G/A                      | 90.9              | 0.023    |
| 8   | rs804276  | NEIL2     | upstream_gene_variant       | G                 | A               | G/A                      | 67.4              | 0.03     |
| 8   | rs8191604 | NEIL2     | intron_variant              | T                 | G               | T/G                      | 69.3              | 0.409    |
| 8   | rs1874546 | NEIL2     | intron_variant              | C                 | G               | C/G                      | 85.2              | 0.769    |
| 8   | rs4639    | NEIL2     | 3_prime_UTR_variant         | A                 | G               | A/G                      | 65.4              | 0.009    |
| 8   | rs1961456 | NAT2      | intron_variant              | A                 | G               | A/G                      | 76.8              | 0.952    |
| 8   | rs1799929 | NAT2      | synonymous_variant          | C                 | T               | C/T                      | 54.3              | 0.374    |
| 8   | rs1799930 | NAT2      | missense_variant            | G                 | A               | G/A                      | 70.9              | 0.17     |
| 8   | rs1799931 | NAT2      | missense_variant            | G                 | A               | G/A                      | 97.3              | 0.814    |
| 8   | rs12674710| NAT2      | downstream_gene_variant     | A                 | C               | A/C                      | 89.8              | 0.538    |
| 8   | rs1495741 | NAT2      | regulatory_region_variant   | G                 | A               | A/G                      | 79.6              | 0.509    |
| 8   | rs13278062| TNFRSF10A | upstream_gene_variant       | G                 | T               | G/T                      | 55.7              | 0.633    |
| 8   | rs1126452 | EPHX2     | synonymous_variant          | A                 | C               | A/C                      | 69.9              | 0.325    |
| 8   | rs3136717 | POLB      | intron_variant              | C                 | T               | T/C                      | 71.6              | 0.439    |
| 8   | rs17226566| LY96      | intron_variant              | T                 | C               | T/C                      | 84.3              | 0.611    |
| 8   | rs9642880 | CASC11    | intron_variant,non_coding_transcript_variant | G | T | T/G | 50.3 | 0.405 |
| 8   | rs2294008 | PSCA      | 5_prime_UTR_variant         | C                 | T               | T/C                      | 54.4              | 0.426    |
| 10  | rs1937845 | AKR1C3    | 5_prime_UTR_variant         | A                 | G               | G/A                      | 52.8              | 0.886    |
| 10  | rs3763676 | AKR1C3    | 5_prime_UTR_variant         | A                 | G               | A/G                      | 72                | 0.289    |
| CHR | SNP    | GENES | VARIATION              | Reference variant | Mutated variant | Major allele/Minor allele | Major.allele.freq | Allelic # |
|-----|--------|-------|------------------------|-------------------|-----------------|--------------------------|------------------|----------|
| 10  | rs12529| AKR1C3| missense_variant       | C                 | G               | G/C                      | 52,8             | 0,886    |
| 10  | rs1937843| AKR1C3| intron_variant         | A                 | G               | A/G                      | 72               | 0,289    |
| 10  | rs4881400| AKR1C3| intron_variant         | T                 | G               | T/G                      | 76,5             | 0,029    |
| 10  | rs12775701| AKR1C3| intron_variant         | A                 | G               | G/A                      | 59,3             | 0,665    |
| 10  | rs2475377| CYP2C9| intergenic_variant     | C                 | T               | C/T                      | 87,3             | 0,283    |
| 10  | rs9332197| CYP2C9| intron_variant         | T                 | C               | T/C                      | 98,2             | 0,132    |
| 10  | rs12357751| BLNK   | intron_variant         | C                 | T               | C/T                      | 73,5             | 0,66     |
| 10  | rs3789928| BLNK   | intron_variant         | G                 | C               | G/C                      | 52,8             | 0,057    |
| 10  | rs2031920| CYP2E1| upstream_gene_variant  | C                 | T               | C/T                      | 97,9             | 0,249    |
| 10  | rs915908| CYP2E1| intron_variant         | G                 | A               | G/A                      | 84,5             | 0,869    |
| 11  | rs7104333| VPS37C| downstream_gene_variant| G                 | A               | G/A                      | 52,5             | 0        |
| 11  | rs625978| GSTP1 | regulatory_region_variant | C       | T               | C/T                      | 61,5             | 0,73     |
| 11  | rs1695  | GSTP1 | missense_variant       | A                 | G               | A/G                      | 62,5             | 0,376    |
| 12  | rs3213427| CD4    | 3_prime_UTR_variant    | T                 | C               | T/C                      | 64,2             | 0,761    |
| 12  | rs11046349| AICDA  | 3_prime_UTR_variant    | T                 | G/A             | T/G                      | 73,6             | 0,283    |
| 12  | rs10878176| TBK1    | intron_variant         | G                 | C               | G/C                      | 71,9             | 0,807    |
| 12  | rs1866074| TDG    | intron_variant         | A                 | G               | A/G                      | 59,4             | 0,247    |
| 12  | rs3990995| UNG    | upstream_gene_variant  | T                 | C               | T/C                      | 81,4             | 0,796    |
| 12  | rs4755621| SCARB1 | intron_variant         | C                 | T               | C/T                      | 51,6             | 0,155    |
| 13  | rs1050112| PARP4  | missense_variant       | G                 | T               | G/T                      | 50,6             | 0,164    |
| 13  | rs13428 | PARP4  | missense_variant       | C                 | G               | C/G                      | 58,8             | 0,131    |
| 13  | rs17655 | ERCC5  | missense_variant       | G                 | C               | G/C                      | 67,7             | 0,887    |
| 14  | rs2228026| TEP1   | synonymous_variant     | A                 | G               | A/G                      | 94,4             | 0,475    |
| 14  | rs1130409| OSGEP  | upstream_gene_variant  | T                 | G               | T/G                      | 53,8             | 0,978    |
| 14  | rs696   | NFKBIA | 3_prime_UTR_variant    | C                 | T               | C/T                      | 52,5             | 0,236    |
| 14  | rs10133290| CHURC1 | downstream_gene_variant| A                 | C               | A/C                      | 72,7             | 0,797    |
| 14  | rs7101  | FOS    | 5_prime_UTR_variant    | C                 | T               | T/C                      | 81,8             | 0,698    |

Chr: Chromosome; SNP: Single Nucleotide Polymorphism; P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval
| CHR | SNP       | GENES   | VARIATION                | Reference variant | Mutated variant | Major allele/Minor allele | Major.allele.freq | Allelic # |
|-----|-----------|---------|--------------------------|-------------------|----------------|--------------------------|-------------------|-----------|
| 14  | rs861539  | KLC1    | intron_variant           | G                 | A              | G/A                      | 60,8              | 0,985     |
| 15  | rs1048943 | CYP1A1  | missense_variant         | T                 | C              | T/C                      | 96,7              | 0,587     |
| 15  | rs2472299 | CYP1A1  | intergenic_variant       | A                 | G              | G/A                      | 58,6              | 0,601     |
| 15  | rs762551  | CYP1A2  | intron_variant           | C                 | A              | A/C                      | 57,6              | 0,901     |
| 15  | rs976072  | POLG    | downstream_gene_variant  | A                 | G              | A/G                      | 71,1              | 0,393     |
| 16  | rs1013358 | MPG     | intron_variant           | C                 | T              | T/C                      | 68,1              | 0,01      |
| 16  | rs1799801 | ERCC4   | synonymous_variant       | T                 | C              | T/C                      | 67,7              | 0,044     |
| 16  | rs17855750| APOBR/IL27 | downstream_gene_variant  | A                 | C              | A/C                      | 96,1              | 0,021     |
| 16  | rs153109  | IL27    | upstream_gene_variant    | T                 | C              | T/C                      | 75,6              | 0,632     |
| 16  | rs1800566 | NQO1    | missense_variant         | G                 | A              | G/A                      | 73,8              | 0,507     |
| 16  | rs689452  | NQO1    | intron_variant           | C                 | G              | G/C                      | 86,1              | 0,61      |
| 16  | rs899729  | IL17C   | upstream_gene_variant    | C                 | A              | C/A                      | 67,3              | 0,598     |
| 17  | rs4791489 | MAP2K4  | downstream_gene_variant  | C                 | T              | C/T                      | 74,7              | 0,251     |
| 17  | rs4796030 | LIG3    | 3_prime_UTR_variant      | A                 | C              | C/A                      | 52,6              | 0,525     |
| 17  | rs2333227 | MPO     | upstream_gene_variant    | C                 | T              | C/T                      | 68,7              | 0,275     |
| 17  | rs7209435 | MAP3K3  | intron_variant           | T                 | C              | T/C                      | 64,5              | 0,09      |
| 17  | rs11655650| BIRC5   | intron_variant           | C                 | T              | C/T                      | 56,9              | 0,485     |
| 18  | rs288980  | ROCK1   | intron_variant           | T                 | C              | C/T                      | 71,4              | 0,03      |
| 18  | rs10775480| SLC14A1 | intron_variant           | T                 | C              | C/T                      | 61                | 0,363     |
| 18  | rs10853535| SLC14A1 | intron_variant           | C                 | T              | T/C                      | 61,3              | 0,479     |
| 18  | rs723279  | SOCS6   | intron_variant           | G                 | A              | G/A                      | 69,1              | 0,963     |
| 19  | rs3746162 | SBN02   | intron_variant           | C                 | T              | C/T                      | 77,5              | 0,392     |
| 19  | rs5498    | ICAM5   | upstream_gene_variant    | A                 | G              | A/G                      | 62,8              | 0,664     |
| 19  | rs25487   | XRCC1   | missense_variant         | T                 | C              | C/T                      | 56,7              | 0,969     |
| 19  | rs1799782 | XRCC1   | missense_variant         | G                 | A              | G/A                      | 94,3              | 0,765     |
| 19  | rs13181   | ERCC2   | downstream_gene_variant  | T                 | G              | T/G                      | 65,9              | 0,092     |
| 20  | rs7265992 | GSS     | intron_variant           | G                 | A              | G/A                      | 90,4              | 0,148     |

Chr: Chromosome; SNP: Single Nucleotide Polymorphism; P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval
### Snp Genotyping

SNP genotyping was performed by NGS on the Ion Torrent Proton sequencer (Life Technologies, Ion Torrent technology). The GeneRead DNAseq Panel PCR Kit V2 (Qiagen) was used for the amplification of the targeted SNPs. The Gene Read primers provided in different three pools. The samples were all re-quantified by Picogreen (using Fluoroscan) and normalized to 20 ng of each sample was plated into 3 daughter plates for the analysis of each group of primers. A validated in-house protocol was used to perform multiplex PCRs in 10 µL reaction volumes, containing 20 ng DNA, 60 nM of each primer pool and 0.73 µL of HotStarTaq enzyme. Amplication was carried out in a 96-well format plates using the standard conditions recommended by Qiagen: 15 min at 95°C and 22 cycles of 15 s at 95°C and 4 min at 60°C and 10 min at 72°C. Amplication products from the same sample were pooled purified using paramagnetic Serapure beads at 1.8X and quantified with the Qubit DNA high-sensitivity assay kit (Invitrogen Corporation). Barcoded libraries were prepared from 10µl of purified products and as previously described [24]. For template preparation, the barcoded libraries were pooled in equimolar concentrations and 7 µL of 100 pM was subsequently used for emulsion PCR (emPCR) using the Ion Torrent Template Ion PI Hi-Q OT2 200 kit on the Ion OneTouch2 Instrument (Thermo Fisher Scientific). The sequencing reaction was performed on an Ion Proton System using Life Technologies' Ion PI™ Chip Kit v3 and Ion PI™ Hi-Q™ Sequencing 200 Kit (Thermoscience Fisher Scientific) (Fig. 3).

### Bioinformatics Analyses

Primary analyses were done using the Torrent_Suite 5.6.0 and included signal processing, base calling, reads alignment to the human genome reference 19 (Genome Reference Consortium GRCh37), quality control of mapping and coverage analysis, and variant calling. Subsequently, a list of detected sequence variants, including SNPs and small insertions/deletions, was imported into (Thermo Fisher Scientific) for annotation. The GeneRead data analysis workflow runs the GATK Variant Annotator program using the TVC output to populate the INFO field in the VCF file with the parameters necessary for downstream filtering. The INFO field of the GATK Unified Genotyper output already contains all the necessary parameters. Alignments were visually verified with the Integrative Genomics Viewer (IGV) v2.3. A combination of Snpeff v4.3, SnpSift v4.3 were used for annotation and filtration, and PLINK and in-house algorithms (R packages ...) were used for analysis. QC filtered variants were then merged, annotated, and filtered using various databases including, ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/), HGMD (http://www.hgmd.cf.ac.uk/ac/index.php), and dbSNP (https://www.ncbi.nlm.nih.gov/snp/).

### Results

All assays were performed using DNA samples from a total of 544 cases and controls and duplicated samples for quality control. Only 40 samples (40/544; 7.35%) failed the library preparation and were excluded from the analysis. In this series, two runs led to 75% and 88% ISP loading and to the generation of total bases of 4.33 and 11.4 Gigabases for total bases of 33, 329, 31 and 84,548,378 respectively, 99% of which aligned to the reference genome (Hg19) (Fig. 4). This analysis provided high analytical sensitivity and allowed detecting SNPs. Using this approach, we have generated molecular profiles of a large number of individuals and identified specific BCA variations.
Among the 139 retained SNP (Table 2) nine minor alleles were significantly inversely associated with bladder cancer development \([\text{rs162555 (CYP1B1*C)}, \text{rs2228000 (XPC*A)}, \text{rs10936599 (ACTRT3*T)}, \text{rs710521 (TP53*C)}, \text{rs3752645 (PRKAR2B*A)}, \text{rs804276 and rs4639 (NEIL2*A and NEIL2*G)}, \text{rs4881400 (AKR1C3*G) and rs288980 (ROCK1*T)})]. When we used the co-dominant model (the reference group is composed of subjects with homozygous genotype for major allele) we also found a statistically significant inverse association for heterozygous genotypes of \(\text{ACTRT3}}\) (rs10936599), \(\text{TP63}}\) (rs710521), \(\text{CYP3A4}}\) (rs2740574), \(\text{BLNK}}\) (rs3789928) and \(\text{COMT}}\) (rs4680) (OR1 = 0.51, 0.62, 0.61, 0.55, and 0.66 respectively) or homozygous genotypes of \(\text{CYP1B1}}\) (rs162555), \(\text{XPC}}\) (rs2228000), \(\text{NEIL2}}\) (rs4639), \(\text{TP63}}\) (rs710521), \(\text{CYP3A4}}\) (rs2740574), \(\text{BLNK}}\) (rs3789928) and \(\text{COMT}}\) (rs4680) (OR2 = 0.36, 0.12, 0.49, 0.38, 0.4 and 0.42 respectively) (Table 3). However six minor alleles in six genes were significantly associated with increased risk of bladder cancer development: \(\text{MPG}}\) (rs1013358 T > C), \(\text{GSTA4}}\) (rs17614751 G > A), \(\text{GSS}}\) (rs7260770 G > A); \(\text{VPS37C/CD5}}\) (rs7104333 G > A), \(\text{APOBR/IL27}}\) (rs17855750 A > C); and \(\text{ERCC4}}\) (rs1799801 T > C). These aggravating effects were attributed to homozygous genotypes for \(\text{VPS37C}}\) (rs7104333, A/A) or \(\text{MPG}}\) (rs1013358, C/C) genes (OR2 = 2.53 and 2.34 respectively) or to the heterozygous genotype for \(\text{ARNT}}\) gene (rs1889740, rs2228099, rs2256355, rs2864873), \(\text{GSTA4}}\) (rs17614751) and \(\text{APOBR/IL27}}\) (rs17855750) with OR1 respectively estimated at 1.99, 2.00, 2.00, 1.47, 1.96 and 2.27 compared to reference group. When comparing the distribution of these unfavorable genotypes according to exposition to environmental risk factors (Tables 4 and 5), we have found that non-smokers patients harboring heterozygous genotypes for \(\text{ARNT}}\)/rs2864873 A > G, \(\text{ARNT}}\)/rs1889740 (C > T) or \(\text{GSTA4}}\)/rs17614751 (G to A) were respectively at 2.775, 3.069 and 6.608 folds increased risk of Bca development compared to non-smokers controls with wild genotypes. When we analyzed genotypic distribution in heavy smokers (BCa and Controls smoking more than 20PY) we found that smoker's patients harboring \(\text{ARNT}}\) CT (rs1889740), \(\text{ARNT}}\) CG (rs2228099), \(\text{ARNT}}\) TC (rs2864873) or \(\text{ROCK1}}\) TC genotypes were respectively associated with 2.690, 2.864, 2.779 and 3.00 fold-increased risk of Bca. In the other hand we have found that \(\text{ARNT}}\) CT (rs1889740), \(\text{ARNT}}\) CG (rs2228099), \(\text{ARNT}}\) TC (rs2864873) and \(\text{GSS}}\) GA genotypes were associated with an increased risk of Bca even in absence of professional risk factors (comparison of Non-exposed patients to Non-exposed controls). However when we compare genotypic distribution of unfavorable genotypes in exposed patients and controls according to professional risk factor, we observed that \(\text{ARNT}}\) TT (rs1889740), \(\text{ARNT}}\) GG (rs2228099), \(\text{ARNT}}\) CC (rs2864873), \(\text{ROCK1}}\) CC/CT genotypes were respectively associated with 12.000, 12.666, 12.666, 3.976, 3.960 fold-increased risk of Bca development.
Table 3
Comparison of genotypic distribution between case and control groups

| Gene | SNP    | Raisonnement/Alleles | MjAF | p-value | OR (95% CI) | p-value | OR (95% CI) | p-value | OR_{CD1} (95% CI)* |
|------|--------|----------------------|------|---------|-------------|---------|-------------|---------|------------------|
| ARNT | rs1889740 | T C > T C | 0.00 | 1.7     | 1.17        | 2.48    | 0.26        | 0.77    | 0.49 1.21 | 0.00 | 1.99 3.14 | 2.1 |
| ARNT | rs2228099 | G C > G C | 0.00 | 1.72    | 1.18        | 2.5     | 0.26        | 0.77    | 0.49 1.21 | 0.00 | 2.34 3.14 | 2.1 |
| ARNT | rs2256355 | C T > C T | 0.00 | 1.73    | 1.19        | 2.53    | 0.31        | 0.8     | 0.51 1.24 | 0.00 | 2.34 3.14 | 2.1 |
| ARNT | rs2864873 | G A > G A | 0.16 | 1.29    | 0.9        | 1.84    | 0.14        | 0.68    | 0.41 1.14 | 0.05 | 1.47 1.01 | 2.1 |
| CYP1B1| rs162555 | C T > C T | 0.06 | 0.72    | 0.5        | 1.02    | 0.02        | 0.39    | 0.17 0.9  | 0.03 | 0.79 0.55 | 1.1 |
| XPC  | rs2228000 | A G > A G | 0.19 | 0.78    | 0.54        | 1.13    | 0.00        | 0.12    | 0.03 0.53 | 0.00 | 0.92 0.62 | 1.1 |
| ACTR3| rs10936599 | T C > T C | 0.00 | 0.52    | 0.35        | 0.79    | 0.62        | 0.76    | 0.26 2.23 | 0.01 | 0.51 0.34 | 0.1 |
| TP63 | rs710521  | C T > C T | 0.01 | 0.64    | 0.44        | 0.91    | 0.89        | 0.95    | 0.44 2.06 | 0.04 | 0.62 0.42 | 0.1 |
| GSTA4| rs17614751 | A G > A G | 0.02 | 2.11    | 1.1        | 4.04    | 0.24        | NA      | NA    0.03 | 1.96 1.02 | 3.1 |
| CYP3A4| rs2740574 | C C > T T/C | 0.03 | 0.67    | 0.46        | 0.96    | 0.35        | 1.51    | 0.63 3.59 | 0.02 | 0.61 0.42 | 0.1 |
| NEIL2| rs4639  | G A > G A | 0.03 | 0.68    | 0.48        | 0.97    | 0.03        | 0.56    | 0.33 0.94 | 0.03 | 0.76 0.52 | 1.1 |
| NEIL2| rs179930 | A G > A G | 0.86 | 0.97    | 0.68        | 1.37    | 0.00        | 0.36    | 0.18 0.74 | 0.01 | 1.14 0.79 | 1.1 |
| AKR1C3| rs4881400 | G T > G T | 0.11 | 0.75    | 0.52        | 1.07    | 0.02        | 0.42    | 0.2 0.91  | 0.05 | 0.84 0.58 | 1.1 |
| BLNK | rs3789928 | C G > C G | 0.01 | 0.58    | 0.39        | 0.86    | 0.78        | 0.94    | 0.62 1.43 | 0.02 | 0.55 0.36 | 0.1 |
| VPS3C | rs7104333 | A G > A G | 0.01 | 1.73    | 1.17        | 2.55    | 0.00        | 2.03    | 1.34 3.09 | 0.00 | 1.42 0.94 | 2.1 |
| MPG  | rs1013358 | C C > T T/C | 0.09 | 1.36    | 0.95        | 1.92    | 0.00        | 2.2     | 1.26 3.84 | 0.01 | 1.15 0.79 | 1.1 |
| ERCC4| rs1799801 | C T > C T | 0.15 | 1.3     | 0.91        | 1.84    | 0.05        | 1.8     | 1 3.22  | 0.09 | 1.18 0.81 | 1.1 |
| APOBR| rs17855750 | C A > C A | 0.01 | 2.36    | 1.16        | 4.79    | 0.49        | NA      | NA    0.02 | 2.27 1.11 | 4.1 |
| ROCK1| rs288980 | T T > C C/T | 0.21 | 0.8     | 0.56        | 1.13    | 0.01        | 0.43    | 0.23 0.8  | 0.02 | 0.94 0.65 | 1.1 |
| GSS  | rs7260770 | A G > A G | 0.03 | 1.62    | 1.05        | 2.5     | 0.97        | 1.02    | 0.25 4.14 | 0.08 | 1.67 1.07 | 2.1 |
| COMT | rs4633  | T C > T C | 0.03 | 0.67    | 0.46        | 0.96    | 0.99        | 1       | 0.64 1.57 | 0.07 | 0.63 0.42 | 0.1 |
| COMT | rs6680  | A G > A G | 0.01 | 0.63    | 0.44        | 0.92    | 0.83        | 0.95    | 0.61 1.49 | 0.04 | 0.6 0.41 | 0.1 |

P: P value; OR: Odds ratio; 95% CI: Confidence Interval, OR_{CD1}: Heterozygous genotype VS Homozygous Wild genotype; OR_{CD2}: Mutant homozygous genotype genotype; p-value ¥: P value for co-dominant model
Table 4
Comparison of genotypes distribution of significant unfavorable SNP from “Tunisian BCHP” panel in bladder cancer

| Gene/Variant | Genotypes | Controls | Bladder cancer | Bladder cancer (NE) Vs Controls (NE) | Bladder cancer (E) Vs Controls (E) | Bladder cancer (E) Vs Controls (E) |
|--------------|-----------|----------|----------------|--------------------------------------|------------------------------------|------------------------------------|
|              | NE < 20PY | ≥ 20PY   | E              | NE < 20PY | ≥ 20PY | p | OR (95% CI) | p | OR (95% CI) | P    |
| ARNT rs1889740 C > T | CC | 44 | 24 | 34 | 58 | 5 | 8 | 50 | 58 | 1* | . | 1* | . |
|                | CT | 43 | 35 | 23 | 58 | 15 | 19 | 91 | 110 | 0.044 | 3.069 (1.026–9.184) | 0.009 | 1.896 (1.169–3.075) | 0.01 | |
|                | TT | 30 | 11 | 11 | 22 | 7 | 4 | 27 | 31 | 0.254 | 2.053 (0.595–7.080) | 0.305 | 1.409 (0.730–2.716) | 0.45 | |
| ARNT rs2228099 C > G | CC | 43 | 23 | 34 | 57 | 5 | 8 | 48 | 56 | 1* | . | 1* | . |
|                | CG | 44 | 36 | 23 | 59 | 15 | 19 | 93 | 112 | 0.054 | 2.931 (0.979–8.771) | 0.007 | 1.932 (1.189–3.138) | 0.01 | |
|                | GG | 30 | 11 | 11 | 22 | 7 | 4 | 27 | 31 | 0.270 | 2.006 (0.581–6.925) | 0.283 | 1.434 (0.741–2.772) | 0.47 | |
| ARNT rs2256355 T > C | TT | 30 | 22 | 33 | 55 | 5 | 8 | 46 | 54 | 1* | . | 1* | . |
|                | TC | 44 | 36 | 24 | 60 | 15 | 19 | 93 | 112 | 0.207 | 2.045 (0.671–6.228) | 0.010 | 1.901 (1.165–3.100) | 0.23 | |
|                | CC | 43 | 12 | 11 | 23 | 7 | 4 | 29 | 33 | 0.970 | 0.976 (0.283–3.371) | 0.253 | 1.461 (0.761–2.803) | 0.00 | |
| ARNT rs2864873 A > G | AA | 51 | 26 | 37 | 63 | 7 | 10 | 71 | 81 | 1* | . | 1* | . |
|                | AG | 42 | 36 | 23 | 59 | 16 | 17 | 79 | 96 | 0.040 | 2.775 (1.044–7.377) | 0.317 | 1.265 (0.797–2.008) | 0.15 | |
|                | GG | 24 | 8 | 8 | 16 | 4 | 4 | 18 | 22 | 0.773 | 1.214 (0.324–4.549) | 0.855 | 1.069 (0.5189 to 2.204) | 0.11 | |
| CYP1B1 rs162555 (T to C) | TT | 63 | 38 | 36 | 74 | 18 | 16 | 106 | 122 | 1* | . | 1* | . |
|                | TC | 47 | 24 | 27 | 51 | 7 | 12 | 60 | 72 | 0.179 | 0.521 (50.201–1.349) | 0.509 | 0.856 (0.540–1.357) | 0.33 | |
|                | CC | 7 | 8 | 5 | 13 | 2 | 3 | 2 | 5 | 1.000 | 1.000 (0.190–5.240) | 0.007 | 0.233 (0.079–0.680) | 0.09 | |
| XPC rs2228000 (G to A) | GG | 72 | 43 | 42 | 85 | 22 | 20 | 114 | 134 | 1* | . | 1* | . |
|                | GA | 37 | 19 | 20 | 39 | 5 | 11 | 52 | 63 | 0.127 | 0.442 (0.154–1.262) | 0.921 | 1.024 (0.632–1.661) | 0.72 | |

P: P value; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Smoker; E: Smoker
| Gene/Variant | Genotypes | Controls | Bladder cancer | Bladder cancer (NE) Vs Controls (NE) | Bladder cancer (E) Vs Controls (E) | Blad/Cont |
|-------------|-----------|----------|----------------|------------------------------------|------------------------------------|----------|
| TP63, P3H2 rs710521 T/C | AA | 2 | 8 | 6 | 14 | 0 | 0 | 2 | 2 | 0.779 | 0.644 (0.029–13.924) | 0.001 | 0.090 (0.020–0.408) | 0.531 |
| | TT | 55 | 46 | 35 | 81 | 16 | 23 | 107 | 130 | , | 1* | , | 1* | |
| | TC | 56 | 21 | 28 | 49 | 9 | 7 | 53 | 60 | 0.195 | 0.552 (0.225–1.355) | 0.257 | 0.763 (0.477–1.218) | 0.001 |
| | CC | 6 | 3 | 5 | 8 | 2 | 1 | 8 | 9 | 0.874 | 1.145 (0.210–6.237) | 0.482 | 0.701 (0.259–1.890) | 0.401 |
| ACTRT3/MYNN rs10936599 (C to T) | CC | 76 | 51 | 44 | 95 | 18 | 24 | 135 | 159 | , | 1* | , | 1* | |
| | CT | 39 | 18 | 19 | 37 | 9 | 5 | 29 | 34 | 0.954 | 0.974 (0.400–2.369) | 0.026 | 0.549 (0.323–0.933) | 0.001 |
| | TT | 2 | 1 | 5 | 6 | 0 | 2 | 4 | 6 | 0.903 | 0.827 (0.038–17.969) | 0.384 | 0.597 (0.187–1.905) | 0.661 |
| GSTA4 rs17614751 (G to A) | GG | 114 | 63 | 63 | 126 | 23 | 31 | 168 | 199 | , | 1* | , | 1* | |
| | GA | 3 | 7 | 5 | 12 | 4 | 1 | 17 | 18 | 0.017 | 6.608 (1.385–31.530) | 0.894 | 0.949 (0.442–2.038) | 0.051 |
| | AA | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0.431 | 4.872 (0.094–251.797) | 0.457 | 3.170 (0.151–66.582) | 0.491 |
| CYP3A4 rs2740574 C>T | CC | 5 | 3 | 1 | 4 | 3 | 3 | 5 | 8 | , | 1* | , | 1* | |
| | CT | 46 | 23 | 28 | 51 | 2 | 8 | 50 | 58 | 0.010 | 0.072(0.009–0.542) | 0.379 | 0.568 (0.161–2.000) | 0.691 |
| | TT | 66 | 44 | 39 | 83 | 22 | 20 | 113 | 133 | 0.445 | 0.555(0.122–2.516) | 0.724 | 0.801 (0.233–2.744) | 0.691 |
| PRKAR2B rs3752645 (G to A) | GG | 88 | 57 | 58 | 115 | 22 | 25 | 149 | 174 | , | 1* | , | 1* | |
| | GA | 26 | 10 | 10 | 20 | 5 | 6 | 19 | 25 | 0.629 | 0.769 (0.265–2.231) | 0.554 | 0.826 (0.438–1.556) | 0.011 |
| | AA | 2 | 3 | 0 | 3 | 0 | 0 | 0 | 0 | 0.878 | 0.786 (0.036–16.971) | 0.119 | 0.094 (0.004–1.847) | 0.141 |
| NEIL2 rs804276 G/A | GG | 39 | 37 | 30 | 67 | 18 | 17 | 82 | 99 | , | 1* | , | 1* | |
| | GA | 59 | 25 | 31 | 56 | 7 | 12 | 67 | 79 | 0.005 | 0.257 (0.098–0.672) | 0.844 | 0.954 (0.601–1.515) | 0.011 |

P: P value; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Smoker; E: Smoker
| Gene/Variant | Genotypes | Controls | Bladder cancer | Bladder cancer (NE) Vs Controls (NE) | Bladder cancer (E) Vs Controls (E) | Bladder Cont |
|-------------|-----------|----------|----------------|--------------------------------------|----------------------------------|--------------|
|             | AA        | 19 8 7 15 2 2 19 21 | 0.063 0.228 (0.047–1.085) | 0.885 0.947 (0.455–1.969) | 0.02* |
| NEIL2 rs4639 | A/G       | 37 36 29 65 16 15 82 97 | 0.017 0.319 (0.124–0.819) | 0.795 1.064 (0.664–1.703) | 0.01* |
| NAT2 rs179930 | G > A     | 56 38 32 70 11 19 83 102 | 0.329 1.549(0.642–3.738) | 0.630 1.118(0.664–1.703) | 0.84* |
| AKR1C3 rs4881400 | T > G     | 46 26 28 54 14 10 78 88 | 0.071 0.441(0.181–1.075) | 0.00! |
| BLNK rs3789928 | G/C      | 59 37 42 79 17 13 70 83 | 0.678(0.135–3.998) | 0.194 0.639 (0.325–1.256) | 0.00! |
| VPS37C/CD5 rs7104333 | G > A | 53 32 35 67 11 14 76 90 | 0.678(0.135–3.998) | 0.194 0.639 (0.325–1.256) | 0.00! |
| MPG rs1013358 | C/T      | 46 24 19 43 6 6 49 45 | 0.194 0.639 (0.325–1.256) | 0.194 0.639 (0.325–1.256) | 0.00! |
| ERCC4 rs1799801 | T/C | 10 5 6 11 5 4 25 29 | 0.153 0.411(0.121–1.391) | 0.092 0.519(0.241–1.115) | 0.04! |

P: P value; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Smoker; E: Smoker
| Gene/Variant | Genotypes | Controls | Bladder cancer | Bladder cancer (NE) Vs Controls (NE) | Bladder cancer (E) Vs Controls (E) | Bladder cancer (E) Vs Controls (NE) |
|-------------|-----------|----------|----------------|-------------------------------------|-----------------------------------|-------------------------------------|
|             | TT        | 58       | 33             | 34                                  | 67                                | 10                                 | 15                                 | 70                                 | 85                                 | 1*                                 | 1*                                 | 1*                                 | 1*                                 |
|             | TC        | 52       | 29             | 29                                  | 58                                | 13                                 | 16                                 | 71                                 | 87                                 | 0.421                             | 1.450 (0.586–3.585)                | 0.4768                             | 1.182 (0.745–1.875)                | 0.586                              |
|             | CC        | 7        | 8              | 5                                   | 13                                | 4                                  | 0                                  | 27                                 | 27                                 | 0.093                             | 3.314 (0.817–13.438)               | 0.188                             | 1.637 (0.784–3.414)                | 0.034                              |
| APOB       | rs17855750 A/C |         |               |                                     |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |
|             | AA        | 111      | 67             | 65                                  | 132                               | 26                                 | 27                                 | 151                                | 178                                | 1*                                |                                    | 1*                                 |                                    |                                    |
|             | AC        | 6        | 3              | 3                                   | 6                                 | 1                                  | 4                                  | 16                                 | 20                                 | 0.757                             | 0.711 (0.082–6.167)                | 0.059                             | 2.471 (0.965–6.326)                | 0.128                              |
|             | CC        | 0        | 0              | 0                                   | 0                                 | 0                                  | 0                                  | 1                                  | 1                                  | 0.475                             | 4.207 (0.081–216.956)              | 0.624                             | 2.226 (0.090–55.100)               | 0.701                              |
| ROCK1       | rs288980 T/C |         |               |                                     |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |
|             | TT        | 13       | 8              | 14                                  | 22                                | 0                                  | 1                                  | 14                                 | 15                                 | 1*                                |                                    | 1*                                 |                                    |                                    |
|             | TC        | 42       | 28             | 22                                  | 50                                | 12                                 | 10                                 | 66                                 | 76                                 | 0.160                             | 7.941 (0.440–143.219)              | 0.035                             | 2.229 (1.056–4.705)                | 0.288                              |
|             | CC        | 62       | 34             | 32                                  | 66                                | 15                                 | 20                                 | 88                                 | 108                                | 0.195                             | 6.696 (0.377–118.925)              | 0.017                             | 2.400 (1.163–4.951)                | 0.319                              |
| GSS         | rs7260770 G/A |         |               |                                     |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |
|             | GG        | 94       | 61             | 56                                  | 117                               | 17                                 | 24                                 | 129                                | 153                                | 1*                                |                                    | 1*                                 |                                    |                                    |
|             | GA        | 20       | 9              | 11                                  | 20                                | 9                                  | 7                                  | 36                                 | 43                                 | 0.057                             | 2.488 (0.970–6.377)                | 0.094                             | 1.644 (0.918–2.943)                | 0.355                              |
|             | AA        | 3        | 0              | 1                                   | 1                                 | 1                                  | 0                                  | 3                                  | 3                                  | 0.605                             | 1.843 (0.180–18.781)               | 0.474                             | 2.294 (0.235–22.338)               | 0.558                              |
| COMT        | rs4633 C>T |         |               |                                     |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |
|             | CC        | 38       | 19             | 19                                  | 38                                | 8                                  | 9                                  | 75                                 | 84                                 | 1*                                |                                    | 1*                                 |                                    |                                    |
|             | CT        | 60       | 35             | 35                                  | 70                                | 14                                 | 15                                 | 62                                 | 77                                 | 0.833                             | 1.108 (0.424–2.891)                | 0.006                             | 0.497 (0.301–0.821)                | 0.037                              |
|             | TT        | 18       | 16             | 14                                  | 30                                | 5                                  | 7                                  | 31                                 | 38                                 | 0.663                             | 1.319 (0.377–4.606)                | 0.075                             | 0.573 (0.310–1.057)                | 0.895                              |
| COMT        | rs4680 G>A |         |               |                                     |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |                                    |
|             | GG        | 38       | 20             | 18                                  | 38                                | 9                                  | 10                                 | 75                                 | 85                                 | 1*                                |                                    | 1*                                 |                                    |                                    |
|             | GA        | 62       | 33             | 35                                  | 68                                | 14                                 | 13                                 | 63                                 | 76                                 | 0.919                             | 0.953 (0.376–2.415)                | 0.006                             | 0.499 (0.302–0.826)                | 0.020                              |
|             | AA        | 17       | 17             | 15                                  | 32                                | 4                                  | 8                                  | 30                                 | 38                                 | 0.992                             | 0.993 (0.268–3.679)                | 0.040                             | 0.530 (0.289–0.973)                | 0.996                              |

P: P value; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Smoker; E: Smoker
Table 5
Comparison of genotypes distribution of significant unfavorable SNP from "Tunisian BCHP" panel in bladder cancer patients according to professional risk factors

| Gene/Variant | Genotypes | Controls | Bladder cancer | Controls (NE) Vs Bladder cancer (NE) | Bladder cancer (E) Vs Controls (E) | Bladder cancer (E) Vs Controls (NE) |
|--------------|-----------|----------|----------------|---------------------------------------|-----------------------------------|-----------------------------------|
|              | Non Exposed (NE) | Exposed (E) | Non Exposed (NE) | Exposed (E) | p | OR (95% CI) | p | OR (95% CI) | p | OR (95% CI) |
| **ARNT rs1889740 C > T** | | | | | | | | | | |
| CC | 90 | 12 | 36 | 19 | - | 1* | 1.934(1.168–3.203) | 0.010 | 1.671(0.667–4.166) | 0.269 | 2.537(1.374–4.685) |
| CT | 84 | 17 | 65 | 45 | 0.010 | 1.882(0.455–1.710) | 0.022 | 12.000(1.416–101.672) | 0.123 | 1.764(0.856–3.635) |
| TT | 51 | 1 | 18 | 19 | 0.710 | 0.887(0.456–1.725) | 0.020 | 12.666(1.491–107.602) | 0.108 | 1.821(0.876–3.783) |
| **ARNT rs2228099 C > G** | | | | | | | | | | |
| CC | 88 | 12 | 35 | 18 | - | 1* | 1.929(1.163–3.201) | 0.010 | 1.803(0.720–4.166) | 0.207 | 2.615(1.405–4.864) |
| CG | 86 | 17 | 66 | 46 | 0.010 | 1.897(0.456–1.725) | 0.020 | 12.666(1.491–107.602) | 0.108 | 1.821(0.876–3.783) |
| GG | 51 | 1 | 18 | 19 | 0.724 | 0.972(0.507–1.865) | 0.020 | 12.666(1.491–107.602) | 0.108 | 1.821(0.876–3.783) |
| **ARNT rs2256355 T > C** | | | | | | | | | | |
| TT | 86 | 12 | 34 | 18 | - | 1* | 1.889(1.133–3.149) | 0.014 | 1.803(0.720–4.166) | 0.207 | 2.526(1.357–4.701) |
| TC | 87 | 17 | 65 | 46 | 0.014 | 1.929(1.163–3.201) | 0.020 | 12.666(1.491–107.602) | 0.108 | 1.821(0.876–3.783) |
| CC | 52 | 1 | 20 | 19 | 0.933 | 0.972(0.507–1.865) | 0.020 | 12.666(1.491–107.602) | 0.108 | 1.821(0.876–3.783) |
| **ARNT rs2864873 A > G** | | | | | | | | | | |
| AA | 101 | 13 | 43 | 33 | - | 1* | 1.713(1.055–2.780) | 0.029 | 0.960(0.403–2.283) | 0.926 | 1.404(0.813–2.424) |
| AG | 85 | 16 | 62 | 39 | 0.029 | 1.713(1.055–2.780) | 0.926 | 0.960(0.403–2.283) | 0.926 | 1.404(0.813–2.424) |
| GG | 39 | 1 | 14 | 11 | 0.636 | 0.843(0.415–1.710) | 0.180 | 4.333(0.507–37.030) | 0.710 | 0.863(0.397–1.875) |
| **CYP1B1 rs162555 (T to C)** | | | | | | | | | | |
| TT | 120 | 17 | 72 | 58 | - | 1* | 0.786(0.4920–2.57) | 0.315 | 0.814(0.3199–2.072) | 0.666 | 0.581(0.337–1.000) |
| TC | 89 | 9 | 42 | 25 | 0.315 | 0.786(0.4920–2.57) | 0.666 | 0.814(0.3199–2.072) | 0.050 | 0.581(0.337–1.000) |
| CC | 16 | 4 | 5 | 0 | 0.221 | 0.183–1.482 | 0.024 | 0.033(0.0017–0.648) | 0.054 | 0.062(0.003–1.058) |
| **XPC rs2228000 (G to A)** | | | | | | | | | | |
| GG | 146 | 17 | 79 | 58 | - | 1* | 1.059(0.6562–1.712) | 0.811 | 0.915(0.349–2.397) | 0.858 | 0.925(0.533–1.604) |
| GA | 68 | 8 | 39 | 25 | 0.811 | 1.059(0.6562–1.712) | 0.858 | 0.915(0.349–2.397) | 0.858 | 0.925(0.533–1.604) |
| AA | 11 | 5 | 1 | 0 | 0.090 | 0.168(0.021–1.325) | 0.016 | 0.027(0.001–0.516) | 0.126 | 0.108(0.0063–1.877) |

P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Exposed; E: Exposed
| Gene   | SNP          | Controls | Bladder cancer | Controls (NE) Vs Bladder cancer (E) | Bladder cancer (E) Vs Controls (NE) |
|--------|--------------|----------|----------------|-------------------------------------|-------------------------------------|
| TP63   | rs710521 T/C | TT       | 124            | 12                                 | 78                                  |
|        |              | TC       | 89             | 16                                 | 36                                  |
|        |              | CC       | 12             | 2                                  | 5                                   |
| ACTRT3/MYNN | rs10936599 C to T | CC | 150 | 21 | 95 | 65, 1* |
|        |              | CT       | 70             | 6                                  | 21                                 |
|        |              | TT       | 5              | 3                                  | 3                                   |
| GSTA4  | rs17614751 G to A | GG | 213 | 27 | 107 | 77, 1* |
|        |              | GA       | 12             | 3                                  | 11                                 |
|        |              | AA       | 0              | 0                                  | 1                                   |
| CYP3A4 | rs2740574 C > T | CC | 8   | 1   | 5    | 3, 1* |
|        |              | CT       | 85             | 12                                 | 29                                 |
|        |              | TT       | 132            | 17                                 | 85                                 |
| PRKAR2B| rs3752645 G to A | GG | 178 | 25 | 101 | 75, 1* |
|        |              | GA       | 43             | 4                                  | 18                                 |
|        |              | AA       | 4              | 1                                  | 0                                   |
| NEIL2  | rs804276 G/A | GG       | 93             | 13                                 | 64                                  |
|        |              | GA       | 100            | 15                                 | 45                                 |

P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Exposed; E: Exposed
| SNP          | Genotype | Controls | Bladder cancer | Bladder cancer (NE) Vs Controls (NE) | Bladder cancer (E) Vs Controls (E) | Bladder cancer (E) Vs Controls (NE) |
|-------------|----------|----------|----------------|--------------------------------------|--------------------------------------|--------------------------------------|
| NEIL2 rs4639| A/G      |          |                |                                      |                                      |                                      |
|             | AA       | 32       | 2              | 10                                   | 7                                   | 0.046 (0.208–0.988)                  |
|             |          |          |                |                                      |                                      | 0.858 (0.214–6.336)                  |
|             |          |          |                |                                      |                                      | 0.156 (0.212–1.282)                  |
|             | AG       | 90       | 12             | 60                                   | 42                                  | 0.273 (0.475–1.234)                  |
|             |          |          |                |                                      |                                      | 0.335 (0.267–1.566)                  |
|             |          |          |                |                                      |                                      | 0.352 (0.453–1.926)                  |
|             | GG       | 41       | 3              | 11                                   | 7                                   | 0.016 (0.191–0.844)                  |
|             |          |          |                |                                      |                                      | 0.595 (0.149–2.978)                  |
|             |          |          |                |                                      |                                      | 0.025 (0.151–0.883)                  |
| NAT2 rs1799930| G>A      |          |                |                                      |                                      |                                      |
|             | GG       | 111      | 15             | 56                                   | 42                                  | -                                    |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |           |          |                |                                      |                                      | 1*                                  |
|             | GA       | 89       | 11             | 57                                   | 39                                  | 0.311 (0.799–2.015)                  |
|             |          |          |                |                                      |                                      | 0.603 (0.519–3.089)                  |
|             |          |          |                |                                      |                                      | 0.578 (0.690–1.943)                  |
|             | AA       | 25       | 4              | 6                                    | 2                                   | 0.124 (0.184–1.226)                  |
|             |          |          |                |                                      |                                      | 0.060 (0.292–1.076)                  |
|             |          |          |                |                                      |                                      | 0.040 (0.048–0.932)                  |
| AKR1C3 rs4881400| T>G      |          |                |                                      |                                      |                                      |
|             | TT       | 121      | 22             | 68                                   | 58                                  | -                                    |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             | TG       | 85       | 4              | 45                                   | 23                                  | 0.802 (0.590–1.503)                  |
|             |          |          |                |                                      |                                      | 0.191 (0.677–7.025)                  |
|             |          |          |                |                                      |                                      | 0.044 (0.323–0.985)                  |
|             | GG       | 19       | 4              | 6                                    | 2                                   | 0.241 (0.184–1.226)                  |
|             |          |          |                |                                      |                                      | 0.060 (0.292–1.076)                  |
|             |          |          |                |                                      |                                      | 0.046 (0.211–0.049)                  |
| BLNK rs3789928| G/C     |          |                |                                      |                                      |                                      |
|             | GG       | 50       | 8              | 39                                   | 30                                  | -                                    |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             | GC       | 119      | 19             | 56                                   | 32                                  | 0.059 (0.356–1.020)                  |
|             |          |          |                |                                      |                                      | 0.103 (0.171–1.178)                  |
|             |          |          |                |                                      |                                      | 0.008 (0.246–0.814)                  |
|             | CC       | 56       | 3              | 24                                   | 21                                  | 0.064 (0.291–1.037)                  |
|             |          |          |                |                                      |                                      | 0.395 (0.442–7.873)                  |
|             |          |          |                |                                      |                                      | 0.172 (0.318–1.228)                  |
| VPS37C/CD5 rs7104333| G>A      |          |                |                                      |                                      |                                      |
|             | GG       | 40       | 6              | 41                                   | 21                                  | -                                    |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             | GA       | 105      | 15             | 48                                   | 45                                  | 0.004 (0.256–0.775)                  |
|             |          |          |                |                                      |                                      | 0.779 (0.291–2.522)                  |
|             |          |          |                |                                      |                                      | 0.529 (0.433–1.537)                  |
|             | AA       | 80       | 9              | 30                                   | 17                                  | 0.001 (0.199–0.669)                  |
|             |          |          |                |                                      |                                      | 0.319 (0.160–1.818)                  |
|             |          |          |                |                                      |                                      | 0.017 (0.192–0.851)                  |
| MPG rs1013358| C>T     |          |                |                                      |                                      |                                      |
|             | CC       | 17       | 4              | 19                                   | 14                                  | -                                    |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             | CT       | 84       | 17             | 41                                   | 36                                  | 0.031 (0.205–0.927)                  |
|             |          |          |                |                                      |                                      | 0.431 (0.173–2.116)                  |
|             |          |          |                |                                      |                                      | 0.113 (0.232–1.167)                  |
|             | TT       | 124      | 9              | 59                                   | 33                                  | 0.020 (0.206–0.878)                  |
|             |          |          |                |                                      |                                      | 0.945 (0.276–3.974)                  |
|             |          |          |                |                                      |                                      | 0.005 (0.323–0.722)                  |
| ERCC4 rs1799801| T>C     |          |                |                                      |                                      |                                      |
|             | TT       | 109      | 16             | 51                                   | 34                                  | -                                    |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |
|             |          |          |                |                                      |                                      | 1*                                  |

P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Exposed; E: Exposed
|                  | Controls | Bladder cancer | Bladder cancer (NE) Vs Controls (NE) | Bladder cancer (E) Vs Controls (E) | Bladder cancer (E) Vs Controls (NE) |
|------------------|----------|----------------|--------------------------------------|------------------------------------|-------------------------------------|
| TC               | 97       | 13             | 55                                   | 35                                 | 0.422                               |
|                  |          |                | 1.211 (0.7581 to 1.937)               | 0.594 (0.5304 to 3.026)            | 0.600 (0.670–1.996)                |
| CC               | 19       | 1              | 13                                   | 14                                 | 0.339                               |
|                  |          |                | 1.462 (0.670–3.189)                   | 0.080 (0.795–54.560)               | 0.033 (1.071–5.207)                |
| APOBR rs17855750 |          |                |                                      |                                    |                                     |
| A/C              |          |                |                                      |                                    |                                     |
| AA               | 214      | 29             | 105                                  | 76                                 | ,                                   |
|                  |          |                | ,                                   | ,                                  | 1*                                  |
| AC               | 11       | 1              | 14                                   | 6                                  | 0.023                               |
|                  |          |                | 2.593 (1.138–5.910)                   | 0.452 (19.849)                     | 0.413 (0.549–4.296)                |
| CC               | 0        | 0              | 0                                    | 1                                  | 0.723                               |
|                  |          |                | 2.033 (0.040–103.184)                 | 0.929 (0.045–29.208)               | 0.193 (0.339–208.705)              |
| ROCK1 rs288980   |          |                |                                      |                                    |                                     |
| T/C              |          |                |                                      |                                    |                                     |
| TT               | 28       | 7              | 9                                    | 4                                  | ,                                   |
|                  |          |                | ,                                   | ,                                  | 1*                                  |
| TC               | 88       | 4              | 43                                   | 36                                 | 0.325                               |
|                  |          |                | 1.520 (0.659–3.503)                   | 0.020 (1.239–12.756)               | 0.064 (2.863–8.751)                |
| CC               | 109      | 19             | 67                                   | 43                                 | 0.116                               |
|                  |          |                | 1.912 (0.850–4.300)                   | 0.044 (1.035–15.154)               | 0.071 (2.761–9.341)                |
| GSS rs7260770    |          |                |                                      |                                    |                                     |
| G/A              |          |                |                                      |                                    |                                     |
| GG               | 186      | 25             | 89                                   | 60                                 | ,                                   |
|                  |          |                | ,                                   | ,                                  | 1*                                  |
| GA               | 35       | 5              | 28                                   | 23                                 | 0.070                               |
|                  |          |                | 1.671 (0.957–2.919)                   | 0.235 (0.654–5.609)                | 0.020 (1.116–3.716)                |
| AA               | 4        | 0              | 2                                    | 0                                  | 0.960                               |
|                  |          |                | 1.044 (0.187–5.812)                   | 0.667 (0.008–21.329)               | 0.474 (0.018–6.454)                |
| COMT rs4633      |          |                |                                      |                                    |                                     |
| C>T              |          |                |                                      |                                    |                                     |
| CC               | 71       | 6              | 43                                   | 35                                 | -                                   |
|                  |          |                | 1*                                   | -                                  | 1*                                  |
| CT               | 109      | 21             | 52                                   | 31                                 | 0.352                               |
|                  |          |                | 0.787 (0.476–1.302)                   | 0.008 (0.690–0.707)                | 0.057 (0.576–1.018)                |
| TT               | 45       | 3              | 24                                   | 17                                 | 0.689                               |
|                  |          |                | 0.880 (0.472–1.642)                   | 0.969 (0.216–4.363)                | 0.449 (0.765–3.849)                |
| COMT rs4680      |          |                |                                      |                                    |                                     |
| G>A              |          |                |                                      |                                    |                                     |
| GG               | 69       | 7              | 45                                   | 35                                 | -                                   |
|                  |          |                | 1*                                   | -                                  | 1*                                  |
| GA               | 111      | 19             | 50                                   | 31                                 | 0.1489                               |
|                  |          |                | 0.690 (0.417–1.141)                   | 0.027 (0.121–0.880)                | 0.039 (0.550–0.311–0.972)         |
| AA               | 45       | 4              | 24                                   | 17                                 | 0.525                               |
|                  |          |                | 0.817 (0.439–1.522)                   | 0.814 (0.218–3.306)                | 0.402 (0.744–3.73–1.485)          |

P: P value for Fisher Test; OR: Odds ratio; 95% CI: Confidence Interval; NE: Non Exposed; E: Exposed

Finally the decision-tree analysis produced a three major BCa class. All of these classes were characterized by an intensity of tobacco use ≥ 20 pack years (PY) and subdivided according to the genotype of the CYP1A2 C > A variation (rs762551). The first class (58/249) is defined by the presence of CYP1A2 CC or CYP1A2 CA genotype and also defined by others 8 variations: ARNT C > G (rs2228099), CYP1B1 T > C (rs162555), SOD1 T > C (rs2173962), ROCK1 T > C (rs288980), IL10 A > T (rs1800890), LY96 T > C (rs17226566), AICDA T > G (rs11046349) and MAP2K4 C > T (rs4791489). The second class of BCa group (31/249) is only defined by the intensity of tobacco use ≥ 20PY and the inheritance of the homozygous genotype for rs762551 in CYP1A2 (CYP1A2 A/A), XPC GG or AG genotype (rs2228000, G > A), and MAP2K4 CC or CT genotype (rs4791489, C > T). The third class (25/249) of BCa patients was defined by rs762551 (CYP1A2 C/A or C/C genotype) and 6 other variables (ARNT rs2228099 C > G, MPO rs2333227 C > T, LY96 rs17226566 T > C, CCR2 rs391835 G > A, APOBEC3 rs1014971 C > T and CASP3 rs3087455 T > G) (Fig. 5).
Discussion

As all complex diseases, BCAs isn’t one SNP/gene disorder. Rather, many SNPs with small effects may result to the impairment of key pathways involved in their pathophysiology. The identification of such SNP-signatures represents an analytical challenge requiring the appliance of novel comprehensive statistical approaches. To our knowledge, this is the first study on BCAs in Tunisian population analyzing a large number of SNPs with NGS technique. Indeed, in this study, we report the development and validation of the targeted amplification-based NGS panel analyzed on the Ion Torrent Proton Platform. This panel is named BCHP (BCa Heredity Panel) and used for detecting clinically useful genetic variations associated with the development of BCAs in a total of 249 patients and 255 controls. After filtration the BCHP was composed of 139 SNPs in 97 genes. These SNPs are located in many different regions: introns sites (BCL10, GSTM4, GSTM3...), regulatory region (TP63, NAT2, GSTP1, GSTP2), up and downstream regulation genes sequences (GSTA4, NAT2 APOB/IL27, CYP2D6...), splicing sites (CAS9...). The encoded enzymes modulate biological processes (xenobiotic metabolic process, cellular response to xenobiotic stimulus, Base-excision repair, immunity response...) and/or affect molecular function (Damaged DNA binding, Enzyme model).

Xenobiotic metabolism pathway

In this study, we have found that CYP1B1 C/C (rs162555), NAT2 A/A (rs1799930), AKR1C3 G/G (rs4881400) genotypes and heterozygous genotypes for CYP3A4 (rs2740574), COMT (rs4680) and PRKAR2B (rs3752645) were associated with a decreased risk of BCa development compared to reference group. CYP1B1 encodes for the enzyme which is implicated in the NADPH-dependent electron transport and xenobiotic oxidation [25]. Nevertheless, no previous study has presented conclusive results for an association between CYP1B1 rs162555 and BCa development. The CYP3A4 (Cytochrome P450 3A4) enzyme encoded by the CYP3A4 gene is expressed in adult human liver and it is responsible for the oxidative metabolism of many clinically used drugs. The CYP3A4*1B variant (rs2740574 C to T) is associated with the reduction of the expression of the CYP3A4 enzyme which could explain its effect on the development of BCAs [26]. These results were supported by others studies which reported a significant protective effect of homozygous wild genotype in other cancer types such as prostate cancer, breast cancer, leukemia, gastric cancer, colorectal cancer, and ovarian cancer [27].

The N-acetyltransferase 2 (NAT2) gene encodes for an important phase II xenobiotic-metabolizing enzyme frequently present in liver and intestinal mucosa. It catalyzes the reaction of aromatic and heterocyclic amine carcinogens via O-acetylation and N-acetylation [28–30]. In our case–control study we have found an inverse association between NAT2 A/A genotype (rs1799930) and BCa development. This result concord with those of Lei Quan who reported that NAT2 A/A carriers (rs1799930) were at 50% decreased risk of bladder cancer development compared to rapid acetylator[31]. Catechol-O-methyltransferase (COMT) encodes a phase II enzyme mainly liable for the degradation of catecholamines, like dopamine and noradrenaline [32]. It catalyzes the O-methylation of 2-hydroxyestradiol to yield 2-ME2 [32]. Additionally, it is involved within the inactivation of potential carcinogenic compounds that produce inflammation and catechol estrogens and thus might protect DNA from oxidative damage. Rs4680 (Ex4-12 G > A, or val158met) is the most studied COMT single nucleotide polymorphism (SNP). In our cohort we explain the significant inverse association of heterozygous genotype against BCa development by the modification of enzyme activity associated to this genotype. Indeed it has been reported that valine (val) variant enzyme is 3–4 times more active than the methionine (met) variant [33, 34]. Moreover we observed that PRKAR2B A/A genotype (rs3752645) is inversely correlated with BCa risk. The rs3752645 G/A SNP lies within the PRKAR2B gene, which encodes a regulatory subunit for cyclic adenosine 3’, 5’-monophosphate kinase. Our result confirms previous findings of a GWAS study where the authors showed that the rs3752645 SNP had a strong inverse association with BCa risk [35].

In contrast we have found that homozygous minor genotype or heterozygous genotype for rs1889740 C/T, rs2228009 C/G, rs2864873 A/G in AKR1C3 (rs4881400), AKR1C3 G/G (rs4881400), PRKAR2B (rs3752645), ARNT (rs1889740, rs2228009, rs2864873), GSS (rs7260770), GSTA4 (rs17614751), or in DNA repair pathway [NEIL2 (rs4639), XPC (rs2228000), MPG (rs1013358), ERCC4/XPF (rs1799801)] or in other molecular functions such as cell proliferation, immunology response.

DNA repair pathway

When considering the DNA repair pathway we have found that the minor alleles of the rs4639 and rs2228000 SNPs in NEIL2 and XPC genes respectively were associated with a decreased risk of BCa development compared to reference groups carrying major alleles. However the minor alleles observed in the studied population for the rs1013358 and rs1799801 SNPs in respectively MPG and ERCC4/ XPFR genes increased the risk of bladder cancer.

The NEIL2 (Nei-like DNA glycosylase 2) gene encodes for an enzyme implicated in the first step of the base excision repair (BER) mechanism which consists of cleaving oxidatively damaged bases and introducing a DNA strand break via the associated lyase reaction [39]. The rs4639, is located on 3’UTR of NEIL2[40]. In our study, we found that NEIL2 G/G genotype for rs4639 was associated with 0, 49-fold decreased risk of BCAs. This inverse association could be explained by the fact that the minor allele interacts with some specific miRNA, which activates the BER pathway [41]. Moreover, we have found that rs4639 and rs804276 in NEIL2 gene were in linkage disequilibrium (p-value = 0.03). Beside NEIL2, we found that XPC AA genotype (rs2228000) was associated with a decreased risk of BCa compared to reference group. The XPC gene is located on 3p25 and encodes for an enzyme involved in global genome repair. This
enzyme represents the earliest damage detector by initiating the NER pathway and eliminating the DNA damages induced by chemical and environmental exposures such as aromatic amines and UV light [42]. Our result confirms the study of Zhu Y. et al. which demonstrates a significant association between the presence of XPC rs2228000 AA genotype and a decreased risk of BCA [43]. However, the recent study of Dai Y. et al. suggests that XPC AA genotype may be linked to an increased risk of bladder and breast cancer [44]. These conflicting results and differences in risk associations may be explained by the different etiology and mechanisms of BCA in study populations with different ethnic backgrounds [15].

On the other hand, ERCC4/XPF and MPG enzymes represent the headmaster in DNA metabolic process, DNA repair, and cellular response to DNA damage stimulus. ERCC4 (6p13.12) is a NER gene that plays a key role in DNA repair that protects against genetic instability and carcinogenesis [45]. The ERCC4-rs1799801 is a synonymous variation (Ser835Ser). A significant association between the ERCC4-rs1799801 polymorphisms and increased risk of BCA in Tunisian population could be explained by the effect of this variation on the enzyme activity. Indeed a previously genotypetype-phenotype correlation analysis indicated that the ERCC4-rs1799801 rare homozygous C/C genotype carriers had an increased trend of ERCC4 expression levels [46]. In contradiction with our results, the meta-analysis of Shi T.Y. et al. also did not provide statistical evidence for an association between ERCC4 gene and the overall risk of several human cancers, they also report that stratification of origins of patients and exposure to environmental risk factors. MPG, a BER gene, also plays an important role in DNA repair.

**Others cellular pathway**

In addition to xenobiotic metabolic pathway and DNA repair process we are interested in this study to investigate the impact of others variations in gene implicated in cell proliferation pathway and immunity response in the etiology of BCA. Among the analyzed variations in this panel we have found that APOB/IL27 A/C genotype (rs17855750) was associated with 2.27 fold-increased risk of BCA. The APOB (Apolipoprotein B Receptor)/ IL27 (Interleukin 27) encodes for a protein that has been recognized as a pleiotropic cytokine with both pro- and anti-inflammatory properties [48]. This significant association could be explained by the implication of encoded enzyme in the inflammatory process. In the same context, Zhou B. et al. showed a significant association between IL-27 gene polymorphisms and IL-2 plasma level and reported that rs17855750 GG (CC) genotype was associated with increased risk for muscle invasive bladder carcinoma. Moreover we found that rs7104333 in CD5/VPS37C gene. The Vacuolar protein sorting-associated protein 37C (VPS37C) is a Component of the ESCRT-I complex, and a regulator of vesicular trafficking process. This protein may be involved in cell growth and differentiation and it is known as a negative regulator of T- and B-cell receptor signaling. Its expression has been shown to be implicated in T lymphocytes tolerance toward tumor cells [49]. In contrast we have found that genetic variations in BLNK (rs3789928), TP63 (rs710521), ROCK1 (rs288980) and ACTRT3/MYNN (rs10936599) were associated with a decreased risk of bladder cancer. BLNK gene encodes for a cytoplasmic linker or adaptor protein that plays a critical role in B cell development [50]. It plays an important role in the pro-B cell to pre-B cell transition and B-cell apoptosis via BCR signaling pathway and reported to have a tumor-suppressive function in various hematologic malignancies [51]. Moreover the Human Protein Atlas consortium showed that high expression of BLNK had a favorable prognosis value in urothelial cancer [52]. TP63 (3q27-28) has a homolog sequence to TP53 (tumor suppressor) and TP73[53]. This gene encodes for an enzyme implicated in the control of cell cycle and plays an important role in apoptosis. Abnormal expression of TP63 is associated with a loss of urothelial differentiation [54]. In our study the TP63 T/C genotype (rs710521) was reported to be inversely associated with BCA (p-value = 0.035 and an OR = 0.62; C95% 0.42–0.89). Our results seem to agree with those of Kimeny L. et al. and Lehmann M.-L. et al. who confirmed that this variant plays a role in decreasing the risk of BCA in European patients [55, 56]. We have also found that ROCK1 TT genotype (rs288980) was associated with 0.75 fold-decreased risk of BCA. ROCK1 enzyme is a necessary effector kinase downstream of Rho GTPases, a very important pathway involved in cell migration and has been identified as a possible therapeutic target [57]. The Cancer Genome Project identified three non-synonymous mutations within the ROCK1 gene [58] but the effect of these variations in protein activity or expression was not more elucidated. Finally ACTRT3/ MYNN (Actin Related Protein T3; 3q26.2) gene is poorly described in the literature. The minor T allele of rs10936599 in ACTRT3/ MYNN has been described in Genome-wide association study conducted by Figueroa J. et al. in 2014, to be highly associated with decreased risk of BCA [59], which was confirmed by Wang M. et al. [60]. However, the effect of rs10936599 polymorphism on the MYNN activity is not well known and more biological functional studies are needed to draw more concise conclusions regarding the underlying molecular mechanism of this variation.

**Interaction between genetic and environmental risk factors**

When we compare the distribution of unfavorable genotypes [GSS G/A (rs7260770), ARNT A/G (rs2864873), ARNT C/T (rs1889740) ARNT C/G (rs2228099), ARNT T/C (rs2256355), GSTA4 G/A (rs17614751), ERCC4 C/C (rs1799801), MPG C/C (rs1013358), VPS37C A/A (rs7104333), APOB/IL27 A/C (rs17855750)] between exposed bladder cancer cases and controls according to environmental risk factors we found that ARNT C/T (rs1889740), ARNT A/G (rs2864873) and GSTA4 G/A (rs17614751) genotypes were associated with an increased risk of BCA even in the absence of tobacco risk factors, or professional risk factors. Moreover when we compared exposed cases to exposed controls, we have not found a significant additive effect between the majority of unfavorable genotypes and exposure to environmental risk factors (smokers’ ≥ 20PY or exposed to professional risk factors). The absence of significant additive effect between the majority of unfavorable genotypes and BCA risk could be explained by the fact that the targeted SNPs encoding enzymes were not directly implicated in the metabolism of xenobiotics. For example, the proinflammatory role of APOB/IL27 enzyme could explain the role of rs17855750 in BCA development independently to exposition to environmental risk factors. However this additive effect reached more than 12 fold-increased risk of BCA development for ARNT rs1889740 T/T, ARNT rs2228099 G/G and ARNT rs2256355 C/C genotypes in subjects exposed to professional risk factors. This association may be explained by the fact that ARNT gene encodes a protein that binds to ligand-bound aryl hydrocarbon receptor and promotes the expression of genes involved in xenobiotic metabolism [36]. Whether this association is caused by specific bladder carcinogens present in the work environment warrants further investigations. To conclude our analysis, a decision tree was implemented to create a disease prediction model. This tree has allowed us to define the risk groups most genetically likely to develop BCA. According to the established decision tree, we note the importance of both smoking status (≥ 20PY) and the genotype of CYP1A2 (rs762551) in the development of BCA. The decision-tree analysis produced a three major BCA class. The first major class (58/249) is defined by the presence of CYP1A2 CC or CYP1A2 CA genotype and also defined by others 8 variations: ARNT C > G (rs2228099), CYP1B1 T > C (rs162555), SOD1 T > C (rs2173962), ROCK1 T > C (rs288980), IL10 A > T (rs1800890), LY96 T > C (rs17226566), AICDA T > G (rs11046349) and
MAP2K4 C>T (rs4791489). The second major class of BCa group (31/249) is only defined by the intensity of tobacco use ≥ 20PY and the inheritance of the homozygous genotype for rs762551 in CYP1A2 (CYP1A2 A/A), XPC GG or AG genotype (rs2228000, G > A), and MAP2K4 CC or CT genotype (rs4791489, C > T). This decision tree confirms the crucial role of tobacco consumption in the etiology of BCa. Indeed BCa is considered as a smoking-related cancer [61]. This risk was attributed to many compounds of tobacco such as 4-Aminobiphenyl, 3-Amino-1,4-dimethyl-SHpyrido [4,3-b] indole (Trp-P-1), Toluene, Benzo[a]pyrene, Benzene... [62]. In the other hand CYP1A2 is an enzyme responsible for the metabolism of caffeine and some tobacco compounds. Rs762551 variation in CYP1A2 gene encodes for the CYP1A2*1F allele. The baseline activity of the enzyme is similar in CYP1A2*1F allele carriers and non-carriers. Moreover it has been reported that the presence of rs762551 (A) codes for the "high inducibility" form the CYP1A2 enzyme, characterized by higher enzyme activity in the presence of an inducer such as smoking or heavy coffee consumption[63]. To explain our result we have reanalyzed the CYP1A2 C > A genotype distribution between cases and controls and according to tobacco status. As result we have found that the inheritance of CYP1A2 (CC) or CYP1 A2 (CA) genotype were respectively associated with 6.89 and 9.04 –fold increased risk of BCa in only heavy smokers patients (≥ 20PY) compared to heavy smokers controls. However we have no data about coffee consumption.

Conclusion

We have conducted the first study in Tunisian population to evaluate systematically the association between genetic variations in BCa and environmental factors. We also determined the effect of studied pathway SNPs in comparison with environmental exposition. Once validated, these findings may provide urologists additional genetic information that may help for clinical assessment and treatment decisions. Nevertheless, the underlying mechanisms through which these genes or SNPs affect the clinical behavior of BCas require further studies. Future investigations in our populations and detailed functional characterization are needed to establish predictive or prognostic markers for BCa.

Abbreviations

ACTRT3
Actin Related Protein T3; 3q26.2
APOBR
Apolipoprotein B Receptor
ARNT
Aryl hydrocarbon receptor nuclear translocator
BCa
Bladder cancer
BER
base excision repair
BLNK
B cell linker
COMT
Catechol-O-methyltransferase
CYP1B1
cytochrome P450 family 1 subfamily B member 1
CYP3A4
Cytochrome P450 3A4
dbSNP
database of single nucleotide polymorphisms
DNA
deoxyribonucleic acid
DRC
DNA Repair Capacity
DSBR
double-strand break repair
EDTA
ethylene diamine tetra-acetic acid
ERCC4
Excision Repair Cross-Complementation group 4
GSTM1
Glutathione-s-transferase M1
GWAS
Genome-Wide Association Study
GSS
Glutathione Synthase
GSTA4
Glutathione S-transferase Alpha 4
HG
High-Grade
IARC
International Agency for Research on Cancer
IL27
Interleukin 27
KEGG
Kyoto Encyclopedia of Genes and Genomes
LG
Low-Grade
MIBC
muscle-invasive BCa
MMR
mismatch repair
MPG
N-methylpurine DNA glycosylase
NEIL2
Nei-like DNA glycosylase 2
NAT2
N-acetyltransferase 2
NER
nucleotide excision repair
NGS
Next-Generation Sequencing
NMIBC
non-muscle-invasive BCa
PRKAR2B
Protein Kinase cAMP-dependent type II regulatory subunit beta
ROCK1
Rho associated coiled-coil containing protein kinase 1
SNP
nucleotide polymorphisms
TCR
transcription-coupled repair
TP63
Tumor Protein p63
VPS37C
vacuolar protein sorting-associated protein 37C
WHO
World Health Organization
XPC
Xeroderma Pigmentosum, Complementation group C
XPF
excision repair cross-complementing rodent repair deficiency, complementation group 4

Declarations

Ethics approval and consent to participate

- This trial protocol and recruitment were approved and carried out after the agreement of the ethics committee of the Charles Nicolle Hospital and approved by an Ethics Committee (IEC Project No. 17-35) and an MTA (MATERIAL TRANSFER AGREEMENT MTA/ 2017 / IMP / GCS)/ 0356) from IARC.

- Informed consent was obtained from all participants (in both languages: Arabic/French) prior to enrollment and participation in this study.

- All samples are coded and no patient names appear in the study. All data whether clinical or personal remains anonymous and secret.

Consent for publication

Not applicable. This study did not use identifying images and any clinical and personal details despite written informed consent for publication of their clinical details and/or clinical

Availability of data and material

The dataset is available upon reasonable request to the corresponding author
Competing interests

The authors declare that they have no competing interests

Funding

"No funding was obtained for this study"

Authors’ contributions

IH: Collecting clinical samples and data, designed the panel, molecular analysis, drafted the manuscript and bioinformatics analyzes. SB, HD and FH: Help in Bioinformatics and statistical analyzes. GD and CV: Technical assistance for NGS analyzes. HA: Clinical characterization of the studied population. SZ, ZN and MA: help in collecting samples. MC and JM: designed the study and revised the manuscript. FL: drafted and revised the manuscript. SO: designed the study, performed the statistical analysis and drafted and revised the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgements

The authors thank all patients for participating in this study. We greatly thank also the Urology department, Charles Nicolle hospital, Tunis, Tunisia, Biochemistry Department, Charles Nicolle Hospital, Tunis and the International Agency for Research on Cancer IARC; Lyon, France.

References

1. Antoni S, Ferlay J, Soerjomataram I, et al (2017) Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends. European Urology 71:96–108. https://doi.org/10.1016/j.eururo.2016.06.010
2. Wong MCS, Fung FDH, Leung C, et al (2018) The global epidemiology of bladder cancer: a joinpoint regression analysis of its incidence and mortality trends and projection. Scientific Reports 8:. https://doi.org/10.1038/s41598-018-19199-z
3. Cancer today. http://gco.iarc.fr/today/home. Accessed 15 Jun 2020
4. Bray F, Ferlay J, Soerjomataram I, et al (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians 68:394–424. https://doi.org/10.3322/caac.21492
5. Cassell A, Yunusa B, Jalloh M, et al (2019) Non-Muscle Invasive Bladder Cancer: A Review of the Current Trend in Africa. World J Oncol 10:123–131. https://doi.org/10.14740/wjon1210
6. Burger M, Catto JWF, Dalbagni G, et al (2013) Epidemiology and risk factors of urothelial bladder cancer. Eur Urol 63:234–241. https://doi.org/10.1016/j.eururo.2012.07.033
7. Colin P, Koenig P, Ouzzane A, et al (2009) Environmental factors involved in carcinogenesis of urothelial cell carcinomas of the upper urinary tract. BJU International 104:1436–1440. https://doi.org/10.1111/j.1464-410X.2009.08838.x
8. Skipper PL, Kim MY, Sun H-LP, et al (2010) Monocyclic aromatic amines as potential human carcinogens: old is new again. Carcinogenesis 31:50–58. https://doi.org/10.1093/carcin/bgp267
9. Rouissi K, Ouehmani S, Hamrita B, et al (2011) Smoking and Polymorphisms in Xenobiotic Metabolism and DNA Repair Genes are Additive Risk Factors Affecting Bladder Cancer in Northern Tunisia. Pathol Oncol Res 17:879. https://doi.org/10.1007/s12253-011-9398-3
10. Matic M, Pekmezovic T, Djukic T, et al (2013) GSTA1, GSTM1, GSTP1, and GSTT1 polymorphisms and susceptibility to smoking-related bladder cancer: A case-control study. Urologic Oncology: Seminars and Original Investigations 31:1184–1192. https://doi.org/10.1016/j.urolonc.2011.08.005
11. García-Closas M, Malats N, Silverman D, et al (2005) NAT2 slow acetylation and GSTM1 null genotypes increase bladder cancer risk: results from the Spanish Bladder Cancer Study and meta-analyses. Lancet 366:649–659. https://doi.org/10.1016/S0140-6736(05)67137-1
12. Kumar N, Moreno NC, Feltes BC, et al (2020) Cooperation and interplay between base and nucleotide excision repair pathways: From DNA lesions to proteins. Genetics and Molecular Biology 43:. https://doi.org/10.1590/1678-4685-gmb-2019-0104
13. Greenman C, Stephens P, Smith R, et al (2007) Patterns of somatic mutation in human cancer genomes. Nature 446:153–158. https://doi.org/10.1038/nature05610
14. Pleasance ED, Cheetham RK, Stephens PJ, et al (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 463:191–196. https://doi.org/10.1038/nature08658
15. He J, Shi TY, Zhu ML, et al (2013) Associations of Lys939Gln and Ala499Val polymorphisms of the XPC gene with cancer susceptibility: A meta-analysis. International Journal of Cancer 133:1765–1775. https://doi.org/10.1002/ijc.28089
16. Rouissi K, Bahria IB, Bougatef K, et al (2011) The effect of tobacco, XPC, ERCC2 and ERCC5 genetic variants in bladder cancer development. BMC Cancer 11:101. https://doi.org/10.1186/1471-2407-11-101
17. Leibovici D, Grossman HB, Dinney CP, et al (2005) Polymorphisms in Inflammation Genes and Bladder Cancer: From Initiation to Recurrence, Progression, and Survival. JCO 23:5746–5756. https://doi.org/10.1200/JCO.2005.01.598
18. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27
19. Kanehisa M, Sato Y, Furumichi M, et al (2019) New approach for understanding genome variations in KEGG. Nucleic Acids Res 47:D590–D595. https://doi.org/10.1093/nar/gky962
20. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6323939/. Accessed 30 Sep 2020

21. Ashburner M, Ball CA, Blake JA, et al (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29. https://doi.org/10.1038/75556

22. Jassal B, Matthews L, Viteri G, et al (2020) The reactome pathway knowledgebase. Nucleic Acids Res 48:D498–D503. https://doi.org/10.1093/nar/gkz1031

23. Selinsksi S (2017) Discovering urinary bladder cancer risk variants: Status quo after almost ten years of genome-wide association studies. EXCLI J 16:1288–1296. https://doi.org/10.17179/excli-2017-1000

24. Calvez-Kelm FL, Foll M, Wozniak MB, et al (2016) KRAS mutations in blood circulating cell-free DNA: a pancreatic cancer case-control. Oncotarget 7:78827–78840. https://doi.org/10.18632/oncotarget.12386

25. Wang M, Wang M, Zhang W, et al (2009) Common genetic variants on 8q24 contribute to susceptibility to bladder cancer in a Chinese population. Carcinogenesis 30:991–996. https://doi.org/10.1093/carcin/bgp091

26. Dally H, Edler L, J7?ger B, et al (2003) The CYP3A4*1B allele increases risk for small cell lung cancer: effect of gender and smoking dose. Pharmacogenetics 13:607–618. https://doi.org/10.1097/00008571-200301000-00004

27. Božina N, Bradamante V, Lovrić M (2009) Genetic Polymorphism of Metabolic Enzymes P450 (CYP) as a Susceptibility Factor for Drug Response, Toxicity, and Cancer Risk. Archives of Industrial Hygiene and Toxicology 60:217–242. https://doi.org/10.2478/10004-1254-60-2009-1885

28. Windmill KF, Gaedigk A, Hall PM, et al (2000) Localization of N-acetyltransferases NAT1 and NAT2 in human tissues. Toxicol Sci 54:19–29. https://doi.org/10.1093/toxsci/54.1.19

29. Yamada S, Tang M, Richardson K, et al (2009) Genetic variations of NAT2 and CYP2E1 and isoniazid hepatotoxicity in a diverse population. Pharmacogenomics 10:1433–1445. https://doi.org/10.2217/pgs.09.66

30. Pande JN, Pande A, Singh SPN (2003) Acetylator status, drug metabolism and disease. Natl Med J India 16:24–26

31. Quan L, Chattopadhyay K, Nelson HH, et al (2016) Differential association for N-acetyltransferase 2 genotype and phenotype with bladder cancer risk in Chinese population. Oncotarget 7:40012–40024. https://doi.org/10.18632/oncotarget.9475

32. Chen Y, Yu X, Li T, et al (2016) Significant Association of Catechol-O-Methyltransferase Val158Met Polymorphism with Bladder Cancer Instead of Prostate and Kidney Cancer. Int J Biol Markers 31:110–117. https://doi.org/10.5301/jbm.5000204

33. Chen J, Lipska BK, Halim N, et al (2004) Functional Analysis of Genetic Variation in Catechol-O-Methyltransferase (COMT): Effects on mRNA, Protein, and Enzyme Activity in Postmortem Human Brain. Am J Hum Genet 75:807–821

34. Lachman HM, Papolos DF, Saito T, et al (1996) Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. Pharmacogenetics 6:243–250. https://doi.org/10.1097/00008571-199606000-00007

35. Figueroa JD, Han SS, Garcia-Closas M, et al (2014) Genome-wide interaction study of smoking and bladder cancer risk. Carcinogenesis 35:1737–1744. https://doi.org/10.1093/carcin/bgu064

36. ARNT Gene - GeneCards | ARNT Protein | ARNT Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=ARNT. Accessed 14 Feb 2021

37. Ke HL, Lin J, Ye Y, et al (2015) Genetic Variations in Glutathione Pathway Genes Predict Cancer Recurrence in Patients Treated with Transurethral Resection and Bacillus Calmette–Guérin Instillation for Non-muscle Invasive Bladder Cancer. Ann Surg Oncol 22:4104–4110.

38. Hazra TK, Kow YW, Hatahet Z, et al (2002) Identification and Characterization of a Novel Human DNA Glycosylase for Repair of Cytosine-derived Lesions. J Biol Chem 277:30417–30420. https://doi.org/10.1074/jbc.C200355200

39. Wei H, Kamat A, Chen M, et al (2012) Association of Polymorphisms in Oxidative Stress Genes with Clinical Outcomes for Bladder Cancer Treated with Bacillus Calmette-Guérin. PLoS One 7:. https://doi.org/10.1371/journal.pone.0038533

40. Lachman HM, Papolos DF, Saito T, et al (1996) Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. Pharmacogenetics 6:243–250. https://doi.org/10.1097/00008571-199606000-00007

41. Ateş NA, Tamer L, Ateş C, et al (2005) Glutathione S-Transferase M1, T1, P1 Genotypes and Risk for Development of Colorectal Cancer. Biochem Genet 43:149–163. https://doi.org/10.1007/s10528-005-1508-z

42. Hazra TK, Kow YW, Hatahet Z, et al (2002) Identification and Characterization of a Novel Human DNA Glycosylase for Repair of Cytosine-derived Lesions. J Biol Chem 277:30417–30420. https://doi.org/10.1074/jbc.C200355200

43. Wei H, Kamat A, Chen M, et al (2012) Association of Polymorphisms in Oxidative Stress Genes with Clinical Outcomes for Bladder Cancer Treated with Bacillus Calmette-Guérin. PLoS One 7:. https://doi.org/10.1371/journal.pone.0038533

44. Kang L, Zou X, Zhang G, et al (2019) A variant in a microRNA binding site in NEIL2 3’UTR confers susceptibility to age-related cataracts. FASEB j 33:10469–10476. https://doi.org/10.1096/fj.201802291R

45. Sugasawa K, Ng JMY, Masutani C, et al (1998) Xeroderma Pigmentosum Group C Protein Complex Is the Initiator of Global Genome Nucleotide Excision Repair. Molecular Cell 2:223–232. https://doi.org/10.1016/S1096-3767(00)80132-X

46. Zhu Y, Lai M, Yang H, et al (2007) Genotypes, haplotypes and diplotypes of XPC and risk of bladder cancer. Carcinogenesis 28:698–703. https://doi.org/10.1093/carcin/bgl201

47. Shi TY, He J, Qiu L-X, et al (2012) Association between XPF Polymorphisms and Cancer Risk: A Meta-Analysis. PLoS ONE 7:e38606. https://doi.org/10.1371/journal.pone.0038606

48. Association between XPF Polymorphisms and Cancer Risk: A Meta-Analysis. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3388076/. Accessed 10 Mar 2020
48. Xu X-P, Hua L-Y, Chao H-L, et al (2017) Genetic association between IL-27 rs153109 polymorphism and cancer risk in Chinese population: a meta-analysis. Journal of Receptors and Signal Transduction 37:335–340. https://doi.org/10.3109/10799893.2014.986743

49. Masson-Lecomte A, Maturana EL de, Goddard ME, et al (2016) Inflammatory-Related Genetic Variants in Non–Muscle-Invasive Bladder Cancer Prognosis: A Multimarker Bayesian Assessment. Cancer Epidemiol Biomarkers Prev 25:1144–1150. https://doi.org/10.1158/1055-9965.EPI-15-0894

50. Y M, J R, E C-S, et al (1999) An essential role for BLNK in human B cell development. Science 286:1954–1957. https://doi.org/10.1126/science.286.5446.1954

51. Nakayama J, Yamamoto M, Hayashi K, et al (2009) BLNK suppresses pre-B-cell leukemogenesis through inhibition of JAK3. Blood 113:1483–1492. https://doi.org/10.1182/blood-2008-07-166355

52. Expression of BLNK in cancer - Summary - The Human Protein Atlas. https://www.proteinatlas.org/ENSG00000095585-BLNK/pathology. Accessed 13 Jul 2020

53. Bergholz J, Xiao Z-X (2012) Role of p63 in Development, Tumorigenesis and Cancer Progression. Cancer Microenviron 5:311–322. https://doi.org/10.1007/s12307-012-0116-9

54. Urist MJ, Di Como CJ, Lu M-L, et al (2002) Loss of p63 Expression Is Associated with Tumor Progression in Bladder Cancer. The American Journal of Pathology 161:1199–1206. https://doi.org/10.1016/S0002-9440(10)64396-9

55. Kiemeney LA, Thorlacius S, Sulem P, et al (2008) Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet 40:1307–1312. https://doi.org/10.1038/ng.229

56. Lehmann M-L, Selinski S, Blaszkewicz M, et al (2010) Rs710521[A] on chromosome 3q28 close to TP63 is associated with increased urinary bladder cancer risk. Arch Toxicol 84:967–978. https://doi.org/10.1007/s00204-010-0617-6

57. Yang S, Zhao Y, Tian Y, et al (2018) Common variants of ROCKs and the risk of hypertension, and stroke: Two case-control studies and a follow-up study in Chinese Han population. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 1864:778–783. https://doi.org/10.1016/j.bbadis.2017.12.007

58. Lochhead PA, Wickman G, Mezna M, Olson MF (2010) Activating ROCK1 somatic mutations in human cancer. Oncogene 29:2591–2598. https://doi.org/10.1038/onc.2010.3

59. Figueroa JD, Ye Y, Siddiq A, et al (2014) Genome-wide association study identifies multiple loci associated with bladder cancer risk. Hum Mol Genet 23:1387–1398. https://doi.org/10.1093/hmg/ddt519

60. Wang M, Chu H, Lv Q, et al (2014) Cumulative effect of genome-wide association study-identified genetic variants for bladder cancer. Int J Cancer 135:2653–2660. https://doi.org/10.1002/ijc.28898

61. Simonis K, Shariat SF, Rink M, Urothelial Cancer Working Group of the Young Academic Urologists (YAU) Working Party of the European Association of Urology (EAU) (2014) Smoking and smoking cessation effects on oncological outcomes in nonmuscle invasive bladder cancer. Curr Opin Urol 24:492–499. https://doi.org/10.1097/MOU.000000000000073

62. Talhout R, Schulz T, Florek E, et al (2011) Hazardous Compounds in Tobacco Smoke. Int J Environ Res Public Health 8:613–628. https://doi.org/10.3390/ijerph8020613

63. Wang L, Hu Z, Deng X, et al (2013) Association between Common CYPIA2 Polymorphisms and Theophylline Metabolism in Non-smoking Healthy Volunteers. Basic & Clinical Pharmacology & Toxicology 112:257–263. https://doi.org/10.1111/bcpt.12038

**Figures**
Figure 1

Overview of data analysis workflow
Figure 2

BCHP protein–protein interaction networks using a STRING server a: Overall BCHP panel b: Xenobiotic metabolic process c: Base-excision repair d: Response to drug
Figure 3

Proton experiment workflow
Figure 4

Run2 Report for Auto user CIRC-PROTON-121-TERT Imen2 100418 299
Figure 5
Decision Tree of predictive profiles of Patients and Controls V: Variable

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFile.docx