

1 **Association of the rs1042522 SNP with prostate cancer risk: a study of cancer tissues,**
2 **primary tumor cultures and serum samples from a European Caucasian population.**

3

4 Emily Toscano-Guerra^{1,2,3*}, Valentina Maggio¹, Javier García¹, Maria Eugenia Semidey⁴,
5 Ana Celma⁵, Juan Morote⁵, Inés de Torres⁴, Marina Giralt², Roser Ferrer², Rosanna
6 Paciucci^{1,2*}

7 ¹ Cell Signaling and Cancer Progression Laboratory, Vall d'Hebron Institute of Research
8 (VHIR), Barcelona, Spain.

9 ² Biochemistry Service, Vall d'Hebron Hospital, Barcelona, Spain.

10 ³Universidad Peruana Cayetano Heredia, Lima, Perú.

11 ⁴Department of Pathology, Vall d'Hebron Hospital, Barcelona, Spain

12 ⁵Department of Urology, Vall d'Hebron Hospital, Barcelona, Spain

13

14 *Correspondence: Dr. Rosanna Paciucci, Biochemistry Service, Vall d'Hebron Hospital,
15 Barcelona, Spain. rospacbar@gmail.com; Dr. Emily Toscano-Guerra, Cell Signaling and
16 Cancer Progression Laboratory, Vall d'Hebron Institute of Research (VHIR), Barcelona,
17 Spain. ema.tg27@gmail.com

18

19

20

1 ABSTRACT

2 Prostate cancer (PCa) is the most common cancer in men and the third leading cause of
3 cancer death in Europe. The *TP53* gene, the most frequently mutated gene in human
4 cancer, is a tumor suppressor gene with crucial functions in preventing tumor
5 development. The single nucleotide polymorphism rs1042522, characterized by the
6 substitution of a proline (PRO) for an arginine (ARG) at the position 72 of the p53 protein
7 (P72R SNP), was studied in 12 primary tumor cultures from prostate biopsies of untreated
8 hormone-naïve patients (hnPCs) with aggressive-metastatic cancer (Gleason ≥ 8), 11
9 radical prostatectomies, and a cohort of 94 serums from patient with aggressive prostate
10 cancer using DNA sequencing and melting curve analysis. The results identified a high
11 frequency of the P72R SNP in prostate cancer samples compared to the general European
12 (non-cancer) population, suggesting a very significant association ($p < 0.0001$) between
13 this SNP and the risk of prostate cancer with an odds ratio of 7.937 (IC 95% 5.37-11.00).
14 The G allele (R72) was more frequent in patients with high Gleason scores (≥ 8)
15 suggesting its association to more undifferentiated-malignant PCa lesions.

16

17 INTRODUCTION

18 Prostate cancer (PCa), highly prevalent in developed countries, is the most frequent
19 cancer in men and the third cause of cancer death, after lung and colorectal, in Europe^{1,2}.
20 In the majority of the cases, the cancer is efficiently eradicated and cured by surgical
21 prostatectomy but late-diagnosed or metastatic aggressive cancers show a five-year
22 relative survival rate below 30%³. The tumor suppressor gene *TP53* plays a key role in
23 preventing tumor development. Not surprisingly it has been described as the most
24 commonly mutated gene in human cancers also in PCa⁴. Somatic *TP53* mutations occur
25 more frequently as a later event, and are driver of the aggressive and metastatic cancer,
26 including PCa⁵.

27 Several hereditary genetic variations, found in approximately 5-20% of different cancer
28 types, are associated with pathogenic variants of genes such *BRCA2* and *ATM*. In PCa,
29 pathogenic germline variants of the *TP53* gene have been described with a relative risk
30 of 4.7-8.6, which is comparable with the frequencies described in those genes⁶.

1 Germline variants alleles known as a single-nucleotide polymorphisms (SNPs) are
2 present in >1% of the population, occur naturally, and are considered not to have a severe
3 impact on people health. However, many of these variants may impact the structure or the
4 function of the protein.

5 To study the characteristics of the aggressive metastatic prostate tumors that lead to
6 incurable fatal disease, we established and partially characterized 12 primary tumor
7 cultures derived from prostate biopsies of untreated hormone-naïve patients (hnPCs) with
8 aggressive metastatic cancer. To identify potential predictive markers, the hnPCs cultures
9 were screened for the presence of mutations in driver genes by DNA sequencing. Here,
10 we describe the identification of the *TP53* gene variant allele rs1042522 in the majority
11 of the hnPCs analyzed. The presence of this SNP was confirmed in a cohort of 11 radical
12 prostatectomy tissues and 94 serum samples from Caucasian patients with prostate cancer.
13 The association of the variant allele with prostate cancer risk was investigated using
14 comparisons with the healthy (non-cancer) European (non-Finnish) population from
15 GnomAD v2.1.1.

16 MATERIALS AND METHODS

17 **Prostate primary cultures (hnPCs) from hormone-naïve patients**

18 Primary cultures from prostate tumor needle biopsies of hormone-naïve patients (i.e.,
19 without previous treatments) were established in our laboratory from a cohort of patients
20 selected for (i) high levels of serum PSA (>50 ng/mL), (ii) positive digital rectal
21 examination (DRE), and (iii) Gleason \geq 8. hnPCs were cultured at 37°C in atmosphere of
22 5% CO₂, with complete DMEM-F12 medium, containing: 2 mM of L-Glutamine, 100U
23 of penicillin/mL, 100 μ g streptomycin/mL, 0.1 mM of non-essential amino acids, 1 mM
24 of sodium pyruvate and 7% fetal bovine serum, with Supplement 1X, human FGF-10 (10
25 ng/ μ L), human EGF (20 ng/ μ L), and Vitamin A and E (200 ng/ μ L). Growth factors and
26 vitamins were added freshly whereas supplements were prepared and stored at -20°C until
27 its used. Supplement (100X) was prepared in DMEM-F12 containing glucose (6 mg/mL),
28 transferrin (1000 μ g/mL), human insulin (2.500 μ g/mL), putrescine (97 μ g/mL), sodium
29 selenite (30 μ M) and hydrocortisone (100 μ M).

30 **Radical prostatectomies (RPs) samples**

1 Eleven radical prostatectomies (RPs) samples from patients with low-grade operable
2 tumors (Gleason ≤ 7) were used to compare with hnPCs. Samples were collected by the
3 Service of Urology of the Hospital Vall d'Hebron (HVH), preserved in Optimal Cutting
4 Temperature compound (OCT) and stored at -80°C until frozen sectioning.

5 **Serum samples from PCa patients**

6 A cohort of 94 serum samples from patients with diagnosis of prostate cancer were
7 selected with the criteria: PSA values ≥ 50 ng/mL, males southern European. The study
8 was performed with surplus serum samples from routinely tested patients at the
9 Biochemistry Service of HVH, following protocols reviewed and approved by the HVH
10 Institutional Review Board (Medical Research Ethics Committee, protocol number
11 PR(AG) 96/2015).

12 Serum samples from blood collected in SST yellow tubes containing separator gels, were
13 centrifugated at 3,500 rpm for 15 minutes at 4°C . The resulting serum was collected in 2
14 mL tubes, properly labeled and codified in the biobank system, were then stored at -80°C .
15 Sample processing was performed according to the scientific and ethical guidelines
16 approved by biomedical research law (Decret 1716/2011).

17 **DNA extraction**

18 The genomic DNA from hnPCs and RP was extracted using the DNeasy mini kit (Qiagen)
19 following the manufacturer's guidelines. For the frozen tissue sections from RPs, samples
20 in criotubes were previously mixed with 180 μL of ATL buffer, then homogenized by
21 passing it through a 20-gauge needle attached to an RNase-free syringe about 5-10 times.
22 Successively, samples were incubated with Proteinase K, at 56°C (1 hour) vortexing every
23 20 minutes during the incubation period. For genomic DNA extraction from serum
24 samples, the MagNA Pure 24 Instrument, and total Isolation Kite (Roche) were used, a
25 fully automated system based on using magnetic glass particles technology (MGP), that
26 allows the processing of up to 24 samples in about 70 minutes.

27 **DNA sequencing**

28 The presence of the variant rs104252 of the *TP53* gene was first identified in hnPCs and
29 RPs by analysis of the complementary DNA (cDNA) using RNA reverse transcription
30 (RT) and Sanger sequencing. cDNA synthesis was done with RNA (1 μg) previously
31 verified for good quality ($\text{RIN} \geq 7$), using the NZY M-MuLV kit (NZY Tech) and the

1 2720 Thermocycler (Applied Biosystems), then the cDNAs were sequenced using Sanger
2 technology.

3 **DNA genotyping**

4 To analyze the Pro72Arg SNP of *TP53* genotyping was conducted for all samples using
5 the Melting Curve analysis with a customized Light SNIp assay (TIB MOLBIOL) on a
6 LightCycler 480 II instrument (Roche). Reactions were carried out in a final volume of
7 20 μ L, comprising 2.0 μ L of LightCycler FastStart DNA Master HybProbe mix (Roche),
8 1.0 μ L of LightSNIp mix, 3 mM MgCl₂, and 50 ng of DNA. The ThermoCycler was set
9 up under the conditions outlined in **Table 1**. Subsequently, the melting curves were
10 assessed using the Melt Curve Genotyping software (Roche).

11 **Control group**

12 To study the association of the SNP P72R with prostate cancer, the frequencies found in
13 the (non-cancer) European (non-Finnish) population from GnomAD v2.1.1, containing
14 134 187 samples, were used as a control group.

15 **Gleason score and grade groups**

16 The Gleason score is a classification system that assesses the aggressiveness of prostate
17 tumors based on the architecture of the cancerous tissues and the degree of cells
18 differentiation as observed under a microscope. This score, which ranges from 6 to 10, is
19 the sum of two histological patterns (the most common plus the highest grade) within a
20 single tissue sample (**Table 2**). To better understand and describe the cancer
21 aggressiveness, the Gleason score is adjusted into a grade group encompassing values
22 from 1 to 5. The correlation between the score and its respective group, is shown in **Table**
23 **2**.

24 **Statistics**

25 Analysis of associations of frequencies and proportions were done using Contingency
26 tables and Fisher's test. Values were considered significant when $p \leq 0.05$. These analyses
27 were conducted using GraphPad Prism 9 software. Data and image analyses were
28 performed using Circos tool (<http://circos.ca/>).

29 **RESULTS**

30 **The *TP53* gene variant allele rs1042522**

1 More than 20 different SNPs have been described in the *TP53* gene⁷, including the variant
2 rs1042522, or P72R, in the amino acid 72 of p53 (**Figure 1A**). This variant contains a
3 guanine at position 357 in codon 72 (**CGC**), a modification that results in a change of the
4 encoded amino acid from proline (CCC) to arginine (R or ARG). (**Figure 1B**). **Figure 1C**
5 shows that the arginine induces a decrease in hydrophobicity (turquoise bars) relative to
6 the most abundant proline (P or PRO), which may impact the interaction of the p53
7 protein with other ligands (CCAR2 or HRMT1L2).

8 **P72R and allele frequencies in primary prostate tumor cultures and cancer tissues**

9 To study the potential role of *TP53* mutations in the aggressiveness of unresectable
10 prostate tumors, we examined 12 primary tumor cultures (hnPCs) by cDNA sequencing
11 and identified a mutation at position 357 that changed a cytosine to a guanine. This change
12 is not a somatic mutation but corresponds to the allelic variant rs1042522. The frequency
13 found by cDNA sequencing for the ARG variant was 0.82 (9/11) compared to the PRO
14 variant frequency of 0.18 (2/11) (**Table 3**). To confirm these results, the 12 primary
15 cultures plus 11 prostatectomy tissue samples from different prostate cancer patients were
16 genotyped using gDNA melting curve analysis (**Table 3, Figure 2A**). In hnPCs, 2
17 homozygotes for cytosine (C, proline), 6 homozygotes for guanine (G, arginine), and 4
18 heterozygotes that predominantly expressed ARG were found. In the RP tissues, 4
19 homozygotes for cytosine, 5 homozygotes for guanine, and 4 heterozygotes expressed
20 predominantly ARG. The total frequencies of C and G in primary tumor cultures were
21 0.33 and 0.67, respectively, while in RP tissues the frequency of C was 0.45 and of G
22 0.55, indicating that primary tumor cultures have a higher incidence of the SNP P72R
23 compared to tumor tissues from prostatectomies. These results would suggest that
24 prostate cancers derived from patients with inoperable tumors may differ from cancers
25 that can be cured by surgery.

26 **P72R SNP is associated with prostate cancer risk.**

27 To validate the above results, the serum of 94 patients with prostate cancer (PSA \geq 50
28 ng/ml) were genotyped (**Figure 2A**). P72R is a controversial SNP, found at different
29 frequencies across populations (worldwide), although the majority of sub-populations
30 show lower frequencies (**Figure 2B**). In the selected European (non-Finnish) non-cancer
31 population the frequency of the G allele was 26.3%, while the frequency of the C allele
32 was 73.6% (**Figure 2B**). In the serum samples, the frequency of homozygotes for ARG
33 (hARG) was 53.38% (53/94), homozygotes for PRO (hPRO) were 8.58% (8/94) and

1 heterozygotes was 35.10% (33/94) (**Figure 2C**). The allele frequencies observed in our
2 cohort of serum, hnPCs, and prostatectomy tissues were then compared with the (non-
3 cancer) European (non-Finnish) control population (**Figure 2D**). The results show
4 significant differences in the recurrence rate of the P72R SNP in all samples compared to
5 the control population, suggesting an association of the P72R SNP with prostate cancer.
6 The significance of this association assessed by a contingency analysis (Fisher's test),
7 showed a very significant association between the P72R SNP and prostate cancer in all
8 samples analyzed, with a frequency of 73.93% for the G allele and an odds ratio of 7.937
9 (IC 95% 5.37-11.00) for serum samples ($p < 0.0001$) (shown in **Table 4**).

10 **P72R SNP is not significantly associated with high Gleason tumors.**

11 The ability of a mutant p53 protein to neutralize apoptosis and transform cells in
12 cooperation with EJ-Ras is enhanced when codon 72 encodes ARG (R72)⁸. This R72
13 variant was reported to be able to bind more efficiently and inactivate PGC-1 α ⁴, a
14 transcriptional target of p53-induced apoptosis in prostate cancer cells, thereby promoting
15 cell invasion and metastasis⁴. These observations support our findings that a higher R72
16 allele frequency is present in more aggressive PCa tumors. To study whether the Gleason
17 score (grade) of all genotyped patients was associated to higher frequency of the R72
18 allele, we used the Gleason grading group (1 to 5 instead of 6 to 10) and divided them
19 into 2 different groups: low - medium aggressive tumors (Gleason grade 1 - 3) and highly
20 aggressive tumors (Gleason grade 4 - 5). Patients were genotyped as hARG, hPRO, and
21 heterozygous (ARG-PRO). Data from the cohort of 94 serum patients were plotted on a
22 Circos plot according to Gleason group. **Figure 3A** shows the distribution of patient's
23 proportions in each Gleason group. Although a marked lower proportion of the P72 allele
24 and a remarkably higher frequency of the R72 allele are visible in high Gleason tumors,
25 these differences did not reach statistical significance ($p\text{-value} = 0.2867$) (**Figure 3A and**
26 **3B**).

27 DISCUSSION

28 The tumor suppressor gene *TP53*, reported mutated with a high frequency (53.3%) in
29 aggressive mCRPC^{9,4}, was the focus of this investigation. The finding that 9 out of 12
30 primary tumor cultures from patients with aggressive metastatic PCa had a p53 protein
31 with arginine (CGC) at codon 72 while only 2 cultures had a proline (CCC), was
32 intriguing. An initial explanation for the enrichment of the R72 variant in hnPCs from

1 cancer tissues was the Loss of Heterozygosity (LOH), particularly studied in exon 4 that
2 includes the codon 72 of the protein. In fact, loss of the Proline (C) allele and preferential
3 retention of the Arginine (G) allele was described in primary tumors and metastatic tissues
4 of squamous cell carcinoma of vulva¹⁰, head and neck¹¹, colorectal tumors¹², and lung
5 cancer¹³. In addition, a recent meta-analysis study using the TCGA pan-cancer database
6 and patients selected for the P72R heterozygosity, reported that 31% (127/409) of
7 heterozygotes had lost the P72 allele in the corresponding tumor tissue¹⁴ indicating that
8 the G allele is preferentially selected for tumor development, possibly due to enhanced
9 p73-induced apoptosis¹⁵ and the modulation of tumor metabolism by regulating PGC-
10 1 α ¹⁶. Although proper LOH screening was not performed in our cancer cohort, the
11 genotyping analysis of serum samples, that showed a similar frequency of the R72 variant
12 compared to hnPCs and cancer tissues from prostatectomy, suggested that it is more likely
13 to be an association with PCa risk.

14 Because of the high frequency of arginine in the genotyped DNA from hnPCs and RPs
15 was significantly different from the European non-cancer population, a larger number of
16 samples (serum) from patients with prostate cancer were examined. The results confirmed
17 the high frequency of the guanine (C to G) at position 357 compared to a control European
18 non-cancer population selected from the GnomAD v2.1, indicating a very significant
19 association ($p < 0.0001$) of this SNP and prostate cancer risk, with odds ratio of 7.937 (IC
20 95% 5.37-11.00). Several reports have described a similar significant association of the
21 R72 SNP with the risk of developing cutaneous melanoma, breast cancer, or prostate
22 cancer¹⁷⁻²⁰. The study on PCa compared 187 Iranian patients with 185 control individual
23 negative for cancer²⁰. Although the frequency of G allele in cancer patients (65.77%) was
24 not as high as we report here (73.93%), the association was statistically significant
25 compared to the control ($p < 0.05$).

26 It is worth noting that several reports have shown controversial results²¹⁻²³ or failed to
27 demonstrate a significant association between the rs1042522 variant and the risk of any
28 type of cancer²⁴⁻²⁸, probably due to limited sample size or selection bias. For example, in
29 a colorectal cancer study in the Northern-European population, no significant association
30 between the presence of the SNP and the risk of cancer was reported²⁹, but the healthy
31 control samples selected had a high frequency of ARG allele (61%) in contrast to the
32 estimated 26% frequency of the (non-cancer) European population from GnomAD v2.1.1
33 (**Figure 2**). Similar to our findings, one study in prostate cancer reported an association

1 between Pro/Pro and a lower risk of prostate cancer³⁰. However, a number of other reports
2 have shown conflicting results^{30–32,27,33–35}. Appropriate representativeness in control
3 samples is necessary for accurate results.

4 Beyond these conflicting findings regarding the ARG frequency at codon 72 of p53, it is
5 crucial to consider the potential impact of this variant on cancer development, although
6 its full understanding is not clear. Studies in mice revealed that the presence of the R72
7 variant shows higher incidence of mammary tumors³⁶. In addition, this variant associated
8 to increased phosphorylation of p53 and enhanced transactivation of *CDKN1A* (p21^{WAF1})
9 in response to starving, with consequent augmented growth arrest but reduced apoptosis,
10 a situation that favors survival³⁷. Furthermore, *TP53* mutations tend to preferentially
11 occur in the R72 allele. In the presence of a mutant *TP53* (e.g. R175H, R273H, and
12 A138V) harboring the R72 SNP, an enhanced capacity for migration, invasion, and
13 metastasis in various cancer cell lines, including PCa, has been reported¹⁶. These
14 observations provide strong evidence that ARG at codon 72 affects the tumor suppression
15 activity of p53.

16 Although no significant association of P72R SNP with the Gleason score was found in
17 our study, the G allele (R72) was more frequent in patients with high Gleason (≥ 8 , group
18 4-5) suggesting a possible association to more undifferentiated-malignant PCa lesions
19 (**Figure 3B**). A few studies also evaluated the association between this polymorphism and
20 tumor grades and found no significant association with clinical stage or Gleason grade in
21 PCa^{33,34,38,39}. One study identified a link with a modest influence of this SNP and shorter
22 biochemical recurrence (BCR) after radical prostatectomy³⁵. Another study in the
23 Japanese population³¹, observed a particularly high frequency of the ARG allele in
24 patients with metastases and those with a high-Gleason score (≥ 8). However, this study
25 analyzed the association using the arginine allele as a reference, meaning they calculated
26 the significance level between the Pro/Pro and Pro/Arg genotype compared to arginine.
27 As expected, they found no significant difference, as proline did not differ significantly
28 from the control group.

29 To validate and extend our findings, further investigation with a larger number of samples
30 and consideration of the family history of prostate cancer would be needed. Nevertheless,
31 our present results suggest that a simple but powerful test (gDNA assessment for the P72R
32 SNP) may be useful for the early identification of patients at risk of aggressive metastatic
33 prostate cancer, allowing clinicians to intervene at an early stage.

1 ***Acknowledgments:*** The authors extend their gratitude to the Biochemistry service from
2 Hospital Vall d'Hebron, Barcelona-Spain, for technical support.

3 ***Author contributions***

4 Contributors ETG, and RP conceived and designed the study. ETG, VM, MES, AC and
5 MG collected clinical samples and data. ETG, VM and JG, performed experiments, and
6 analyzed data. ETG and RP generated figures and tables and wrote the manuscript. RP,
7 JM, IT and RF contributed to data interpretation and critically reviewed the manuscript.
8 All authors approved the final manuscript for submission.

9 **The authors declare no conflicts of interest.**

1 REFERENCES

- 2 1. Ferlay, J. *et al.* Global Cancer Observatory: Cancer Today.
3 <http://gco.iarc.fr/today/home> (2020).
- 4 2. Wang, L. *et al.* Prostate Cancer Incidence and Mortality: Global Status and
5 Temporal Trends in 89 Countries From 2000 to 2019. *Frontiers in Public Health* **10**,
6 (2022).
- 7 3. Prostate Cancer - Statistics. *Cancer.Net* [https://www.cancer.net/cancer-](https://www.cancer.net/cancer-types/prostate-cancer/statistics)
8 [types/prostate-cancer/statistics](https://www.cancer.net/cancer-types/prostate-cancer/statistics) (2023).
- 9 4. McCann, J. J. *et al.* Mutant p53 elicits context-dependent pro-tumorigenic
10 phenotypes. *Oncogene* **41**, 444–458 (2022).
- 11 5. Arora, K. & Barbieri, C. E. Molecular Subtypes of Prostate Cancer. *Curr Oncol Rep*
12 **20**, 58 (2018).
- 13 6. Maxwell, K. N. *et al.* Inherited TP53 Variants and Risk of Prostate Cancer.
14 *European Urology* **81**, 243–250 (2022).
- 15 7. T, B., Jld, P. & Me, M. Common genetic variants in the TP53 pathway and their
16 impact on cancer. *Journal of molecular cell biology* **11**, (2019).
- 17 8. Marin, M. C. *et al.* A common polymorphism acts as an intragenic modifier of
18 mutant p53 behaviour. *Nat Genet* **25**, 47–54 (2000).
- 19 9. Robinson, D. *et al.* Integrative clinical genomics of advanced prostate cancer. *Cell*
20 **161**, 1215–1228 (2015).
- 21 10. Brooks, L. A. *et al.* Preferential retention of codon 72 arginine p53 in squamous cell
22 carcinomas of the vulva occurs in cancers positive and negative for human
23 papillomavirus. *Cancer Res* **60**, 6875–6877 (2000).
- 24 11. Schneider-Stock, R. *et al.* Retention of the arginine allele in codon 72 of the p53
25 gene correlates with poor apoptosis in head and neck cancer. *Am J Pathol* **164**,
26 1233–1241 (2004).
- 27 12. Schneider-Stock, R. *et al.* Selective Loss of Codon 72 Proline p53 and Frequent
28 Mutational Inactivation of the Retained Arginine Allele in Colorectal Cancer.
29 *Neoplasia* **6**, 529–535 (2004).

- 1 13. Papadakis, E. D., Soultziz, N. & Spandidos, D. A. Association of p53 codon 72
2 polymorphism with advanced lung cancer: the Arg allele is preferentially retained in
3 tumours arising in Arg/Pro germline heterozygotes. *Br J Cancer* **87**, 1013–1018
4 (2002).
- 5 14. De Souza, C. *et al.* Effect of the p53 P72R Polymorphism on Mutant TP53 Allele
6 Selection in Human Cancer. *J Natl Cancer Inst* **113**, 1246–1257 (2021).
- 7 15. Marin, M. C. *et al.* A common polymorphism acts as an intragenic modifier of
8 mutant p53 behaviour. *Nat Genet* **25**, 47–54 (2000).
- 9 16. Basu, S. *et al.* Mutant p53 controls tumor metabolism and metastasis by regulating
10 PGC-1 α . *Genes & Development* **32**, 230 (2018).
- 11 17. Shen, H. *et al.* p53 Codon 72 Arg Homozygotes Are Associated with an Increased
12 Risk of Cutaneous Melanoma. *J Invest Dermatol* **121**, 1510–1514 (2003).
- 13 18. Volodko, N. *et al.* TP53 codon 72 Arg/Arg polymorphism is associated with a
14 higher risk for inflammatory bowel disease development. *World J Gastroenterol* **21**,
15 10358–10366 (2015).
- 16 19. Dia, Y. *et al.* Arg Allele of P53 Gene Codon72 is a Risk Factor for Breast Cancer in
17 Senegalese Women. *Journal of Molecular and Genetic Medicine* **11**, (2017).
- 18 20. Doosti, A. & Dehkordi, P. G. The p53 codon 72 polymorphism and association to
19 prostate cancer in Iranian patients. *African Journal of Biotechnology* **10**, 12821–
20 12825 (2011).
- 21 21. Ezzikouri, S. *et al.* The Pro variant of the p53 codon 72 polymorphism is associated
22 with hepatocellular carcinoma in Moroccan population. *Hepatology Research* **37**,
23 748–754 (2007).
- 24 22. Mohana Devi, S. *et al.* Analysis of Genetic Damage and Gene Polymorphism in
25 Hepatocellular Carcinoma (HCC) Patients in a South Indian Population. *Dig Dis Sci*
26 **58**, 759–767 (2013).
- 27 23. Son, M. S. *et al.* Promoter polymorphisms of pri-miR-34b/c are associated with
28 hepatocellular carcinoma. *Gene* **524**, 156–160 (2013).

- 1 24. Tang, N.-P., Wu, Y.-M., Wang, B. & Ma, J. Systematic review and meta-analysis of
2 the association between P53 codon 72 polymorphism and colorectal cancer.
3 *European Journal of Surgical Oncology (EJSO)* **36**, 431–438 (2010).
- 4 25. Weng, Y. *et al.* p53 codon 72 polymorphism and Hematological Cancer Risk: An
5 Update Meta-Analysis. *PLOS ONE* **7**, e45820 (2012).
- 6 26. Ye, J., Li, X.-F., Wang, Y.-D. & Yuan, Y. Arg72Pro Polymorphism of TP53 Gene
7 and the Risk of Skin Cancer: a Meta-Analysis. *PLOS ONE* **8**, e79983 (2013).
- 8 27. Zhang, L. *et al.* Association between p53 Pro72Arg polymorphism and prostate
9 cancer risk: a meta-analysis. *J Biomed Res* **25**, 25–32 (2011).
- 10 28. Xu, T., Xu, Z.-C., Zou, Q., Yu, B. & Huang, X.-E. P53 Arg72Pro polymorphism and
11 bladder cancer risk--meta-analysis evidence for a link in Asians but not Caucasians.
12 *Asian Pac J Cancer Prev* **13**, 2349–2354 (2012).
- 13 29. Krüger, S. *et al.* The p53 codon 72 variation is associated with the age of onset of
14 hereditary non-polyposis colorectal cancer (HNPCC). *Journal of Medical Genetics*
15 **42**, 769–773 (2005).
- 16 30. Henner, W. D. *et al.* Association of codon 72 polymorphism of p53 with lower
17 prostate cancer risk. *Prostate* **49**, 263–266 (2001).
- 18 31. Suzuki, K. *et al.* A p53 codon 72 polymorphism associated with prostate cancer
19 development and progression in Japanese. *J Biomed Sci* **10**, 430–435 (2003).
- 20 32. Li, M. S., Liu, J. L., Wu, Y., Wang, P. & Teng, H. Meta-analysis demonstrates no
21 association between p53 codon 72 polymorphism and prostate cancer risk. *Genet*
22 *Mol Res* **10**, 2924–2933 (2011).
- 23 33. Lu, Y. *et al.* Association of p53 codon 72 polymorphism with prostate cancer: an
24 update meta-analysis. *Tumor Biol.* **35**, 3997–4005 (2014).
- 25 34. Sivoňová, M. K. *et al.* Association of p53 and p21 polymorphisms with prostate
26 cancer. *Biomedical Reports* **3**, 707–714 (2015).
- 27 35. Duncan, A. *et al.* Association of TP53 Single Nucleotide Polymorphisms with
28 Prostate Cancer in a Racially Diverse Cohort of Men. *Biomedicines* **11**, 1404
29 (2023).

- 1 36. Gunaratna, R. T. *et al.* Dynamic role of the codon 72 p53 single-nucleotide
2 polymorphism in mammary tumorigenesis in a humanized mouse model. *Oncogene*
3 **38**, 3535–3550 (2019).
- 4 37. Kung, C.-P., Liu, Q. & Murphy, M. E. The codon 72 polymorphism of p53
5 influences cell fate following nutrient deprivation. *Cancer Biology & Therapy* **18**,
6 484–491 (2017).
- 7 38. Babaei, F. *et al.* The TP53 Codon 72 Polymorphism and Risk of Sporadic Prostate
8 Cancer among Iranian Patients. *Iran J Public Health* **43**, 453–459 (2014).
- 9 39. Aydin, M. *et al.* Lack of evidence of HPV etiology of prostate cancer following
10 radical surgery and higher frequency of the Arg/Pro genotype in Turkish men with
11 prostate cancer. *Int Braz J Urol* **43**, 36–46 (2017).

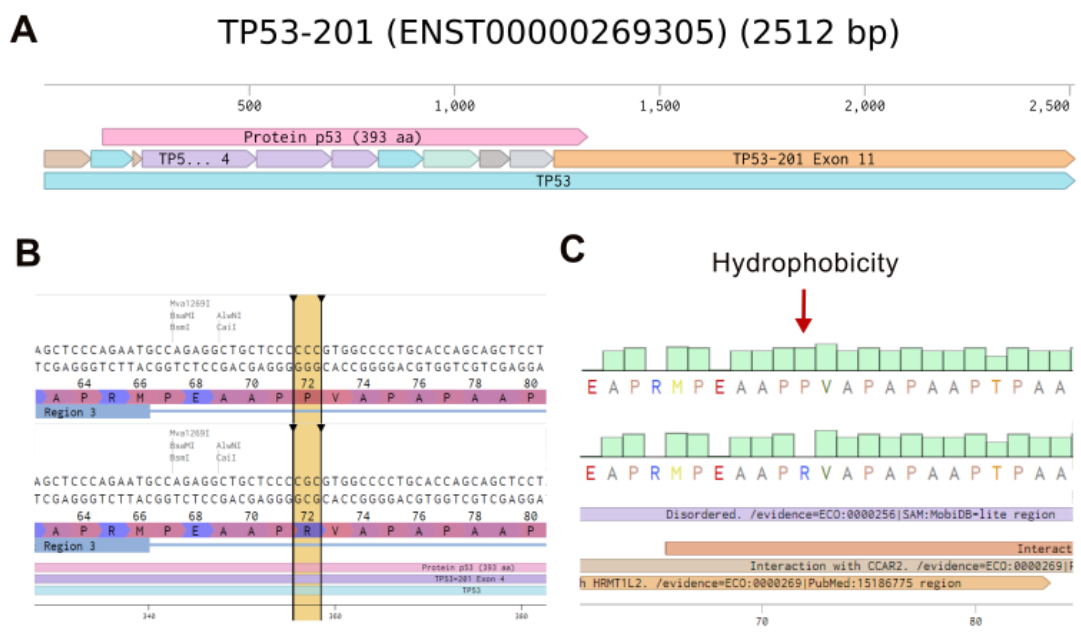
12

13

14

1 **FIGURE LEGENDS**

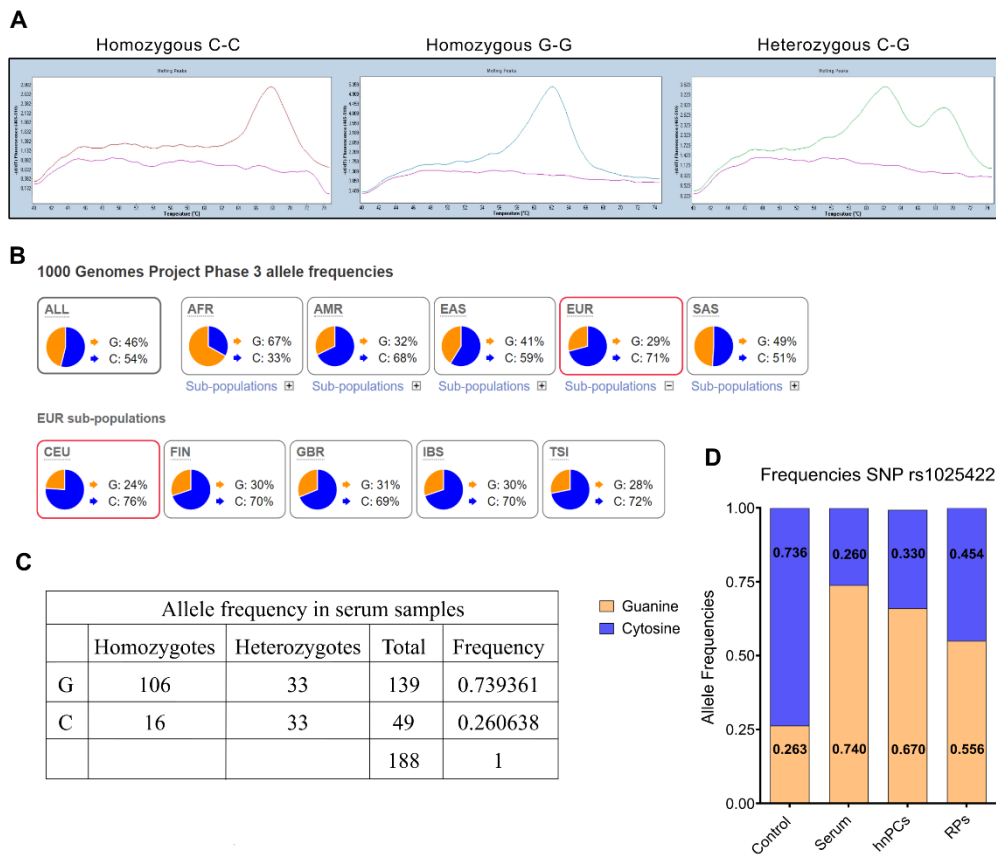
2 **Figure 1. SNP P72R in the TP53 gene. A)** The canonic transcript for variant rs1042522.
 3 The TP53 transcript 201 consists of 11 exons encoding a protein of 393 amino acids
 4 including exons 2 to the 5' region of exon 11. **B)** The SNP P72R (black arrows)
 5 polymorphic variants, Proline (P) encoded by the triplet CCC, and Arginine (R) encoded
 6 by CGC at position 357 of the coding strand. **C)** The change in the hydrophobic profile
 7 of variant rs1042522 (red arrow) when P is changed to R, a polar-positive amino acid.
 8 Figures from Benchling.com.



9

10 **Figure 2: Allele frequencies in the European population and in prostate cancer**
 11 **samples. A)** Representative melting curves from serum samples. Peaks of absorbance at
 12 different dissociation temperatures. Left panel: homozygous samples with cytosine in
 13 both strands show a dissociation peak at 68°C. Center panel: homozygous samples with
 14 guanine in both strands show a dissociation peak at 60°C. Right panel: heterozygous
 15 samples with cytosine and guanine are identified by both peaks. **B)** Global frequency for
 16 the allele rs1042522. Allele G has a lower frequency in all sub populations except for the
 17 African population. Data and Images from Ensemble.org. **C)** rs1042522 frequencies in
 18 serum samples from patients with prostate cancer. **D)** rs1042522 frequencies in primary
 19 tumor cultures (hnPCs), serum samples, radical prostatectomy biopsies, compared to the
 20 European population.

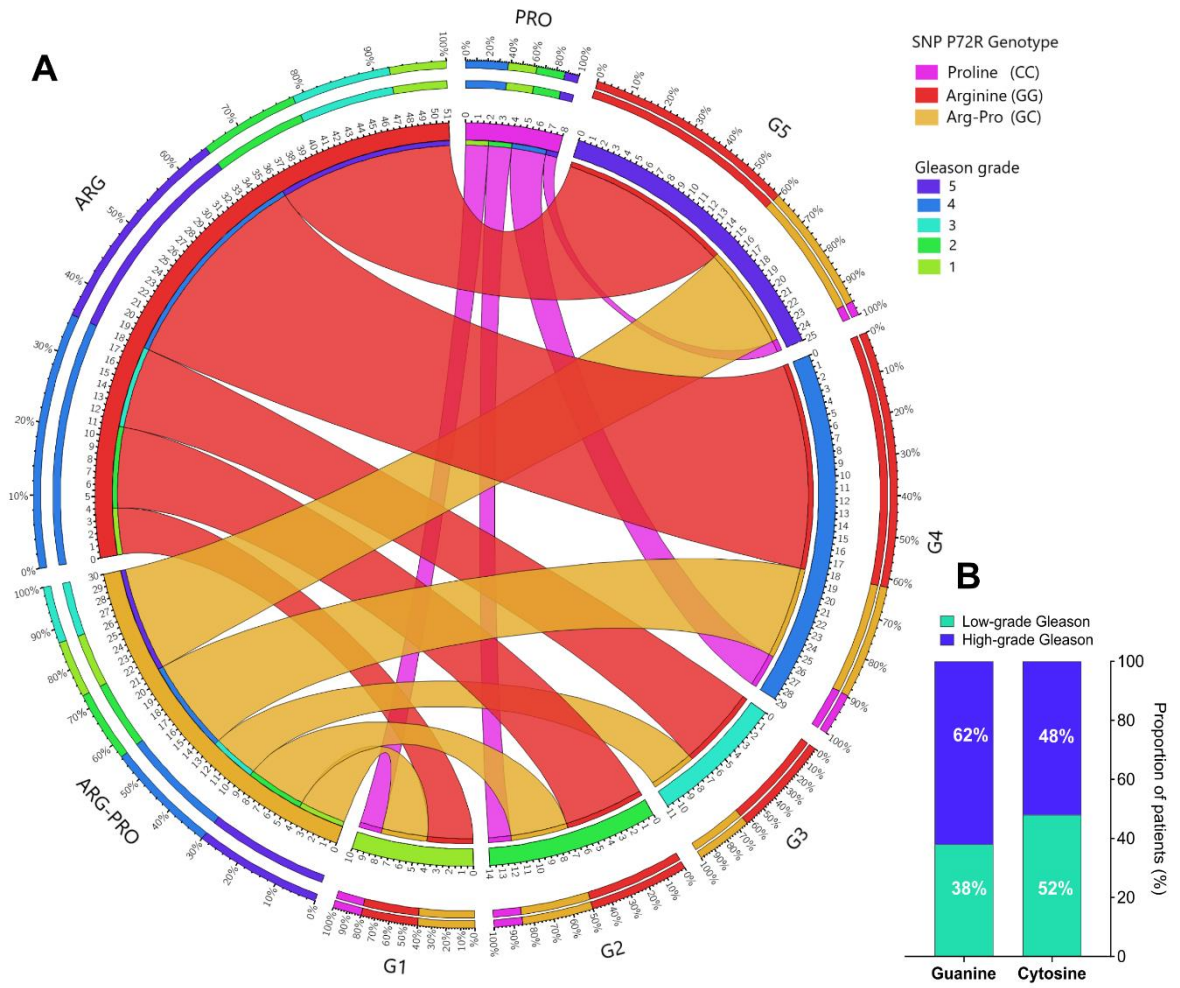
1



2

3 **Figure 3. Distribution of the rs1042522 variant in 94 prostate cancer patients**
 4 **according to Gleason grade. A)** Circos plot displaying the distribution of the P72R
 5 genotypes according to patient Gleason score. **B)** Distribution of the SNP in low Gleason
 6 grade (1-3) and high Gleason grade (4-5); the contingency analysis of these proportions
 7 did not yield significant results. Gleason grade 1: Gleason score ≤ 6 ; Gleason grade 2-3:
 8 Gleason score = 7; Gleason grade 4: Gleason score = 8; Gleason grade 5: Gleason score
 9 ≥ 9 .

1



2

3

4

5

6

7

1 **TABLES:**

2 **Table 1. Cycling conditions for Melting curve analysis**

3

| Phase. | Cycles | Temperature | Hold | Ramp rate (°C/s) |
|---------------------|---------------|--------------------|-------------|-------------------------|
| Denaturation | 1 | 95 | 10 min | 4.6 |
| Cycling | 45 | 95 | 10 sec | 4.6 |
| | | 60 | 10 sec | 2.4 |
| | | 72 | 15 sec | 4.6 |
| Melting | 1 | 95 | 30 sec | 4.6 |
| | | 40 | 2 min | 2.0 |
| | | 75 | 0 sec | - |
| Cooling | 1 | 40 | 30 sec | 2.0 |

4

5

6 **Table 2. Prostate cancer classification in groups according to Gleason pattern.**

| Patterns* | Score | Group | Significance |
|------------------|--------------|--------------|---------------------|
| 1+2, 3+2, 3+3 | ≤ 6 | 1 | Low-grade cancer |
| 3+4 | 7 | 2 | Medium-grade cancer |
| 4+3 | 7 | 3 | Medium-grade cancer |
| 3+5, 4+4, 5+3 | 8 | 4 | High-grade cancer |
| 4+5, 5+4 | 9 | 5 | High-grade cancer |
| 5+5 | 10 | 5 | High-grade cancer |

*Gleason patterns indicate the grade of tissue differentiation.

¹Small, uniform glands, well differentiated.

²More stroma between glands, well differentiated.

³Distinctly infiltrative margins, moderately differentiated.

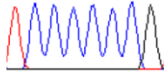
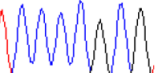
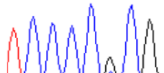
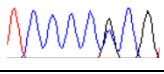
⁴Irregular masses of neoplastic glands, poorly differentiated.

⁵Only occasional gland formation, poorly differentiated/anaplastic tissue.

7

1 **Table 3. P72R genotyping in primary cultures and radical prostatectomy tissues**

2

| SAMPLE | Sanger sequencing | | Expressed Aminoacid | LightSNiP assay | |
|--------|-------------------|---|---------------------|-----------------|---------------------|
| | cDNA | | | gDNA | |
| hnPC01 | CCC | TCCCCCG  | Proline | CC | Homozygous Proline |
| hnPC06 | CCC | | Proline | CC | Homozygous Proline |
| hnPC02 | CGC | | Arginine | GG | Homozygous Arginine |
| hnPC03 | CGC | | Arginine | GG | |
| hnPC08 | CGC | TCCCCGCG  | Arginine | GG | |
| hnPC09 | CGC | | Arginine | GG | |
| hnPC10 | CGC | | Arginine | GG | |
| hnPC12 | CGC | | Arginine | GG | |
| hnPC04 | CGC | TCCCCGCG  | Arginine | CG | |
| hnPC05 | CGC | | Arginine | CG | |
| hnPC11 | CGC | TCCCCGCG  | Arginine | CG | |
| hnPC07 | N.D | N.D | N.D | CG | |
| RP11D | CGC | | Arginine | CG | Heterozygous |
| RP14D | CGC | | Arginine | GG | Homozygous |
| RP17B | CGC | | Arginine | GG | Homozygous |
| RP18F | CGC | | Arginine | GG | Homozygous |
| RP21D | CGC | | Arginine | GG | Homozygous |
| RP23D | CCC | | Proline | CC | Homozygous |
| RP36E | CCC | | Proline | CC | Homozygous |
| RP37A | CGC | | Arginine | CG | Heterozygous |
| RP40B | CCC | | Proline | CC | Homozygous |
| RP63A | CGC | | Arginine | GG | Homozygous |
| RP68C | CCC | | Proline | CC | Homozygous |

1 **Table 4. Analysis of the association of the rs1042522 variant and prostate cancer**

2

| Serum samples | Prostate Cancer, n (%) | GnomAD Non-cancer, n (%) | Odds Ratio (95% CI)* | P- value* |
|--|-------------------------------|---------------------------------|-----------------------------|------------------|
| Arginine (G allele) | 139 (0.74) | 31021 (0.26) | 7.937 (5.73– 11.00) | <0.0001 |
| Proline (C allele) | 49 (0.26) | 86797 (0.74) | | |
| Total | 188 | 117818 | | |
| hnPCs samples | | | | |
| Arginine (G allele) | 16 (0.67) | 31021 (0.26) | 5.596 (2.395– 13.080) | <0.0001 |
| Proline (C allele) | 8 (0.33) | 86797 (0.74) | | |
| Total | 24 | 117818 | | |
| RPs samples | | | | |
| Arginine (G allele) | 12(0.55) | 31021 (0.26) | 3.358(1.451-7.772) | 0.0058 |
| Proline (C allele) | 10(0.45) | 86797 (0.74) | | |
| Total | 22 | 117818 | | |
| Contingency table for the association between the P72R SNP and prostate cancer. The selected European (non-Finnish) non-cancer population from GnomAD was used as a control. *Odds ratio and p-values were calculated with the Fisher's Test. | | | | |

3

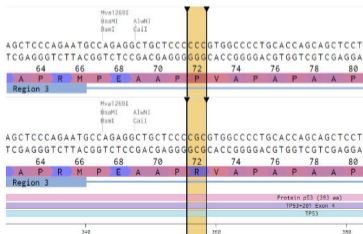
4

5

A TP53-201 (ENST00000269305) (2512 bp)

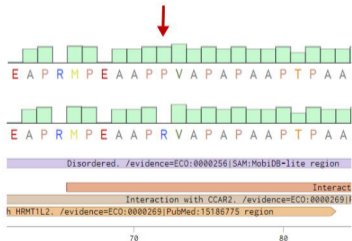


B



C

Hydrophobicity

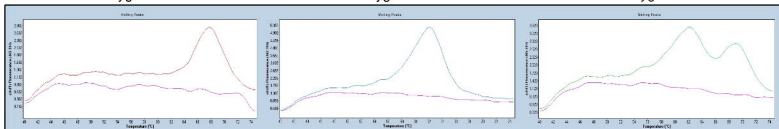


A

Homozygous C-C

Homozygous G-G

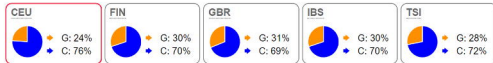
Heterozygous C-G

**B**

1000 Genomes Project Phase 3 allele frequencies



EUR sub-populations

**C**

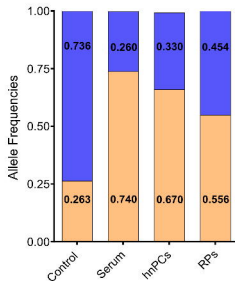
Allele frequency in serum samples

| | Homozygotes | Heterozygotes | Total | Frequency |
|---|-------------|---------------|-------|-----------|
| G | 106 | 33 | 139 | 0.739361 |
| C | 16 | 33 | 49 | 0.260638 |
| | | | 188 | 1 |

■ Guanine
■ Cytosine

D

Frequencies SNP rs1025422



A