Determination of ERG(+), EZH2, NKX3.1, and SPINK-1 subtypes to evaluate their association with clonal origin and disease progression in multifocal prostate cancer

Yenifer Yamile Segura-Moreno 1,2 | María Carolina Sanabria-Salas 1 | Jorge Andrés Mesa-López De Mesa 3 | Rodolfo Varela-Ramirez 4,5 | Natalia Lizeth Acosta-Vega 1 | Martha Lucía Serrano 1,2

Abstract

Background: The prognostic relevance of prostate cancer (PCa) molecular subtypes remains controversial, given the presence of multiple foci with the possibility of different subtypes in the same patient.

Aim: To determine the clonal origin of heterogeneity in PCa and its association with disease progression, SPOP, ERG(+), EZH2, NKX3.1, and SPINK-1 subtypes were analyzed.

Methods: A total of 103 samples from 20 PCa patients were analyzed; foci of adjacent non-tumor prostate tissue, HGPIN, GL3, GL4, GL5, and LN were examined to determine the presence of the TMPRSS2-ERG fusion and ERG, EZH2, NKX3.1, and SPINK-1 expression levels, using RT-PCR. Mutations in exons 6 and 7 of the SPOP gene were determined by sequencing. The presence of subtypes and molecular patterns were identified by combining all subtypes analyzed. To establish the clonal origin of multifocal PCa, molecular concordance between different foci of the same patient was determined. Association of these subtypes with histopathological groups and time to biochemical recurrence (BCR) was assessed.

Results: No mutation was found in SPOP in any sample. The ERG(+) subtype was the most frequent. The molecular pattern containing all four PCa subtypes was only detected in 3 samples (4%), all LN, but it was the most frequent (40%) in patients. Molecular discordance was the predominant status (55%) when all analyzed molecular characteristics were considered. It was possible to find all subtypes, starting as a preneoplastic lesion, and all but one LN molecular subtype were ERG(+) and NKX3.1 subtypes. Only the expression of the NKX3.1 gene was significantly different among the histopathological groups. No association was found between BCR time in patients and molecular subtypes or molecular concordance or between clinicopathological characteristics and molecular subtypes of ERG, EZH2, and SPINK-1.

Conclusion: The predominance of molecular discordance in prostatic foci per patient, which reflects the multifocal origin of PCa foci, highlights the importance of analyzing multiple samples to establish the prognostic and therapeutic relevance of molecular.
subtypes in a patient. All the subtypes analyzed here are of early onset, starting from preneoplastic lesions. NKX3.1 gene expression is the only molecular characteristic that shows a progression pattern by sample.

KEYWORDS
clonal evolution, ERG gene, EZH2 gene, genetic heterogeneity, NKX3.1 gene, prostate cancer, SPINK-1 gene, SPOP gene, TMPRSS2-ERG fusion

1 | BACKGROUND

According to the report of the International Agency for Research on Cancer (IARC), prostate cancer (PCa) is the second type of cancer with the highest incidence and the fifth cause of death by cancer worldwide for both sexes and all ages.\(^1\) PCa starts as a preneoplastic lesion known as high-grade prostatic intraepithelial neoplasia (HGPIN); it progresses to cancer and can cause regional and distant metastasis. In a large percentage of cases, PCa behaves as an indolent disease, but in other cases, it can act aggressively and cause death. The criteria to establish PCa prognosis include Prostate Specific Antigen (PSA) level, Gleason score, and clinical tumor, node, and metastasis (TNM) stage. Nevertheless, these criteria do not establish a prognosis with sufficient certainty.\(^2\)

Numerous studies have proposed different classifications of PCa molecular subtypes that can contribute to determining prognosis; among them, ETS(+) subtypes characterized by chromosomal fusions that cause the overexpression of ETS transcription factors. The most prevalent ETS fusion in PCa is TMPRSS2-ERG, which leads to ERG overexpression.\(^3-5\) ERG overexpression, in turn, generates an increase in the expression of EZH2, which acts as a histone methyltransferase enzyme, and a decrease in the expression of the NKX3.1 tumor suppressor gene, which is an androgen-regulated prostate-specific homeobox gene and a transcription factor. NKX3.1 also binds to TMPRSS2 upstream sequences, and in TMPRSS2-ERG cases, it negatively regulates ERG expression and is a negative feedback control since ERG directly represses NKX3.1.\(^6,7\) EZH2 and NKX3.1 had also been described as PCa subtypes. Another subtype reported in PCa is characterized by SPINK-1 overexpression, which specifically happens in a subset of ETS(−) and is associated with another subtype, the SPOP subtype, characterized by a mutation in this gene.\(^3\) SPINK-1 protein is a trypsin inhibitor, which can function as an autocrine growth factor,\(^8,9\) while SPOP is an E3 ubiquitin ligase substrate-binding subunit of the proteasome complex that mediates the ubiquitination of target proteins, leading more frequently to their proteasomal degradation.\(^10\)

The prognostic relevance of PCa molecular subtypes remains controversial.\(^11-20\) The presence of multiple foci with the possibility of different subtypes in the same patient with PCa could contribute to the difficulty of predicting the clinical behavior of a specific PCa subtype.\(^21-23\) It is not well known whether these foci have a monoclonal origin that expands through the prostate, giving rise to new foci that share some molecular characteristics by origin but begin to acquire changes in subclonal events; or whether each focus appears independently, having a multiclonal origin, so they are expected to be very heterogeneous at the molecular level.\(^22,24,25\)

To elucidate the origin of heterogeneity in PCa and its association with disease progression, ERG(+), SPOP, EZH2, SPINK-1, and NKX3.1 subtypes were analyzed in 103 samples from prostatic foci and regional metastasis to lymph nodes (LN) from 20 PCa patients with poor prognosis. Molecular patterns are also established by sample and patient, combining the presence of all subtypes analyzed.

2 | METHODS

2.1 | Samples

Formalin-fixed paraffin-embedded (FFPE) tissue samples obtained from radical prostatectomy (RP) of 20 PCa patients with poor prognosis (regional metastases and/or biochemical recurrence) and 5-year follow-up were included. Cases were chosen from the project “Exploration of potential predictor biomarkers in patients diagnosed with prostate cancer,” approved by the Ethics Committee of the Instituto Nacional de Cancerología (Bogotá, Colombia) to develop this study. Since tissues were obtained as part of routine care and not specifically for this study, and the data used in this study were previously anonymized, the Ethics Committee of the Instituto Nacional de Cancerología (Bogotá, Colombia) waived the need for informed consent for this study.

Four to five FFPE samples per patient were selected to identify foci with different degrees of differentiation; a total of 83 samples were analyzed. These samples corresponded to HGPIN (n = 17), PCa foci with different Gleason scores (n = 53; GL3 n = 23; GL4 n = 23; and GL5 n = 7), and LN (n = 13). A sample of adjacent non-tumor prostate tissue was used from each patient as a reference (n = 20). Two frozen PCa tissue samples were also used as a positive control of the TMPRSS2-ERG fusion, as well as the PC3 cell line was assessed as a negative control for gene fusions. All methods complied with relevant guidelines and regulations at the national and international levels.

2.2 | Cell culture

The PC3 cell line was derived from androgen-independent PCa. This line was kindly provided by Dr. Fabio Aristizabal, from the Universidad...
Nacional de Colombia. The culture conditions were as follows: 5% CO2 at 37°C, in DMEM medium (Dulbecco/Vogt modified Eagle’s minimal essential medium), supplemented with 10% fetal bovine serum and a cocktail of penicillin, streptomycin, and 1% amphotericin B.

2.3 | DNA/RNA extraction and quantification from FFPE tissue samples

DNA/RNA was isolated from FFPE tissue by automated extraction in the QIAcube using a modified protocol of the commercial kit AllPrep DNA/RNA FFPE (QIAGEN, Hilden, Germany). The modification consisted of incubating the sample before starting the automated protocol with proteinase K at 56°C for 3 h, instead of 15 min. The RNA was quantified by using the Qubit® 2.0 fluorometer RNA HS assay (Invitrogen, Waltham, MA). The RNA was stored at −70°C until its use. Good quality DNA was obtained with this same kit, which was also quantified using the Qubit® 2.0 fluorometer DNA HS assay and stored at −20°C.

2.4 | Analysis of the presence or absence of the TMPRSS2-ERG fusion

RT-PCR was performed to determine the presence of the TMPRSS2-ERG fusion (exon 1 of TMPRSS2: exon 4 of ERG), using the KAPA SYBR FAST One-Step kit to carry out qRT-PCR (Kapa Biosystems, Wilmington, MA), following the recommendations of the manufacturer. The primers used are presented in Table 1.²⁻⁵ It was run in the PTC-200 Thermal Cycler (PCR System) (MJ Research, Reno, NV). The protocol begins with reverse transcription at 42°C for 5 min for 1 cycle; activation of reverse transcriptase at 95°C for 3 min for 1 cycle; denaturation of samples at 95°C for 10 s, annealing at 61°C for 20 s, and extension at 72°C for 1 s for 45 cycles. The presence or absence of the TMPRSS2-ERG fusion was determined by visualizing the corresponding amplimers in 2% agarose gels with SYBR Safe.

2.5 | Analysis of the presence or absence of SPOP mutations

SPOP mutations were identified by sequencing exons 6 and 7, which were amplified through PCR with the Platinum High Fidelity PCR kit (Invitrogen, Waltham, MA), using the SPOP-6 and SPOP-7 primers²⁶ (Table 1), with the following cycling conditions: incubation at 94°C for 2 min, denaturation at 94°C for 30 s, primer annealing at 55°C for 30 s, and extension at 68°C for 1 min; 35 cycles were run. Subsequently, for sequencing, the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystem, Foster City, CA) was used with the same PCR primers, following the manufacturer’s recommendations. The sequences were purified according to the BigDye XTerminator KIT protocol (Applied Biosystem, Foster City, CA), and obtained using the 3500 Genetic Analyzer (Applied Biosystem, Foster City, CA).

### Table 1

| Gene          | Primers                                                                 | Amplicon length (bp) | Annealing temperature (°C) | Primer concentration (nM) | References |
|---------------|-------------------------------------------------------------------------|----------------------|-----------------------------|---------------------------|------------|
| TMPRSS2-ERG   | F: TAGGGCGCGCTAAACGAGGAG R: GTAGGCAACTCAAA CAACGACTGG                  | 125                  | 61                          | 100                       | ⁵          |
| SPOP -6       | F: ACCCATAGCTT TGGTTCTTTCTCCCC R: TATCTGTTTTGGAC AGGTTGTTGCC            | 170                  | 55                          | 200                       | ²⁶         |
| SPOP -7       | F: ACTCATCAGTCTGGGAACCTGC R: AGTTGTGCTTCTGATCTGGT                      | 240                  | 55                          | 200                       | ²⁶         |
| ERG           | F: TCTTGAGACCAACAAAGTGC  R: GTCCGGATCGCTCATCTGG                       | 240                  | 55                          | 200                       | ²⁶         |
| EZH2          | F: TGGTTGATACCTCACCAAGGA R: GAGGACCGTGCTTTTTCAT                        | 90                   | 55                          | 250                       | ²⁸         |
| NKX3.1        | F: CCGAGACGCTGGCAAGAGA R: GAAAGGGCCGCTGAAGTGT                         | 81                   | 55                          | 200                       | ²⁹         |
| SPINK-1       | F: CAAAAATCTGGGCTTGTGAGAAAC  R: AGGCTCCTGGGCTGACCTGT                   | 62                   | 55                          | 200                       | ⁸          |
| UBC           | F: ATTTTTGCGGCGTTGTCTGG  R: TGCCCTGACATTCTCGATG                        | 133                  | 55                          | 150                       | ³⁰         |

Note: F, forward; R, reverse.
2.6 | RT-PCR for ERG, EZH2, NKX3.1, SPINK-1, and UBC gene expression

RT-PCR was carried out in a single step using the qScript 1-Step SYBR Green qRT-PCR kit (Quanta Biosciences, Cummings Center, Beverly, MA), following the manufacturer’s instructions. The protocol begins with reverse transcription at 50°C for 10 min for 1 cycle; activation of reverse transcriptase at 95°C for 5 min for 1 cycle; denaturation of samples at 95°C for 10 s, annealing for 20 s, and extension at 72°C for 1 min for 45 cycles. The primers used for each gene and respective amplicon length, concentration, and annealing temperature are described in Table 1. A LightCycler® 480 thermal cycler (Roche Life Science, Penzberg, Germany) was used for RT-qPCR. To obtain the melting curves, the PCR products were heated at each sample was calculated using the method of Pfaffl, and 2.8 |

| Relative expression of ERG, EZH2, NKX3.1, and SPINK-1 |

With the E values from the standard curves, the relative expression of each sample was calculated using the method of Pfaffl, with the following equation:

$$\text{Relative expression} = \left( \frac{E_{\text{target}}^{CP_{\text{Norm}}} \text{Sample}}{E_{\text{ref}}^{CP_{\text{Norm}}} \text{Sample}} \right) \frac{E_{\text{ref}}^{CP_{\text{Norm}}} \text{Calibrator}}{E_{\text{target}}^{CP_{\text{Norm}}} \text{Calibrator}}$$

where $E_{\text{target}}$ is the E of the target gene; $E_{\text{ref}}$ is the E of the reference gene (UBC); $CP_{\text{target}}$ is the cross point of the target gene; $CP_{\text{ref}}$ is the CP of the reference gene (UBC). Calibrator is the adjacent non-tumor prostate tissue for each patient. To analyze the relative expression of ERG, EZH2, NKX3.1, and SPINK-1, the values obtained were converted into three categories: highly expressed (>1.20), normally expressed (between 0.80 and 1.20), and lowly expressed (<0.80).

2.9 | Determination of PCa molecular subtypes

PCa molecular subtypes were assigned as follows: ERG(-) subtype, when TMPRSS2-ERG fusion is present and/or ERG mRNA is highly expressed; SPOP subtype, in the presence of mutations in the SPOP gene; EZH2 subtype, when EZH2 mRNA is highly expressed; NKX3.1 subtype, when NKX3.1 mRNA is lowly expressed; and SPINK-1 subtype, when SPINK-1 mRNA is highly expressed.

These subtypes were assigned by sample and patient; to consider that a patient is positive for a subtype, at least one sample must have this subtype. A molecular pattern is the combination of the presence of all subtypes analyzed in samples and patients.

2.10 | Concordance patterns of the presence/absence of the TMPRSS2-ERG fusion, SPOP mutations, and gene expression of ERG, EZH2, NKX3.1, and SPINK-1 in prostatic foci of patients

Concordance for the presence/absence of the TMPRSS2-ERG fusion, SPOP mutations, and gene expression of ERG, EZH2, NKX3.1, and SPINK-1 gene expression was assessed separately and together, aiming to establish the clonal origin of all prostatic foci (HGPIN and PCa) analyzed by patient. LN were not used for concordance analyses, given that they are extraprostatic lesions. Figure 1A–E graphically describe the five concordance patterns found: two concordant patterns (patterns A and B), two partially concordant patterns (patterns C and D), and one discordant pattern (pattern E).

2.11 | Data analysis

Clinicopathological characteristics were analyzed as continuous variables: age at diagnosis (years), prediagnosis PSA (ng/ml), tumor percentage in RP, index diameter (cm), and time to BCR (months), as well as categorical variables: Gleason score, Gleason grade group at RP, margin status in RP, perineural invasion, lymphovascular invasion, primary tumor size and/or amount of spread into nearby structures (pT), lymph node compromise (pN), and 5-year BCR. Molecular characteristics were analyzed as categorical variables: detection of the TMPRSS2-ERG fusion, SPOP mutations, and relative expression of ERG, EZH2, NKX3.1, and SPINK-1. The values obtained were converted into three categories: highly expressed (>1.20), normally expressed (between 0.80 and 1.20), and lowly expressed (<0.80). The distribution of the relative expression of ERG, EZH2, NKX3.1, and SPINK-1 was also analyzed.

The assumption of normality in the quantitative variables was validated using the Shapiro–Wilk test. For the variables that fulfilled the
assumption of normality, the mean (SD) was used, while the median (interquartile range) was employed for those that did not. For comparison between groups, Student’s t test for independent samples was applied in the case of variables with normal distribution, and the Wilcoxon signed-rank test otherwise. In the case of categorical variables, absolute and relative frequencies were obtained as descriptive measures. Given the sample size, the Fisher exact test was applied to determine whether there was a statistically significant association between the categorical variables.

The Shapiro–Wilk test showed that relative gene expression data did not comply with the normality criteria in any histopathological group. Although the data complied with normality when using log2 transformation, they did not fulfill the assumption of homogeneity of variances according to Levene’s test. Significant differences were verified by groups with the molecular alterations evaluated in the study, using the Kruskal-Wallis nonparametric test for categorical variables and the Fisher test for quantitative variables. According to the results obtained with the Kruskal-Wallis test, a paired comparison was subsequently applied between histopathological groups and clinicohistopathological characteristics. The results were considered significant with a $p < .05$, using the Mann-Whitney U-test, and visualized according to the distribution of median values and data dispersion, plotted on a logarithmic scale log2 with a box-and-whisker diagram. All analyses were carried out with the statistical software SPSS v18.

3 | RESULTS

3.1 | Clinicopathological characteristics of the cohort

A total of 103 samples from 20 PCa patients were analyzed: foci of adjacent non-tumor prostate tissue ($n = 20$), HGPIN ($n = 17$), GL3 ($n = 23$), GL4 ($n = 23$), GL5 ($n = 7$), and LN ($n = 13$). The main clinicopathological characteristics of the 20 patients are summarized in Table 2.
| Characteristics                          | Full cohort (n = 20) | BCR time* | ≤3 months (n = 9) | >3 months (n = 8) | p-value |
|-----------------------------------------|----------------------|-----------|-------------------|------------------|---------|
| Age at diagnosis (years)                |                      |           |                   |                  |         |
| Median [IQR]                            | 66.0 [12.5]          | 65.0 [12.0]| 63.0 [21.5]       | 1.00             |         |
| Prediagnosis_PSA                        |                      |           |                   |                  |         |
| Median [IQR]                            | 14.3 [12.7]          | 14.3 [9.75]| 11.5 [11.8]       | 0.681            |         |
| Gleason score, n (%)                    |                      |           |                   |                  |         |
| 7                                       | 19                   | 8 (50.0)  | 8 (50.0)          | 1.00             |         |
| 8                                       | 1                    | 1 (100)   | 0 (0.00)          |                  |         |
| Gleason grade group at RP (%)           |                      |           |                   |                  |         |
| G2                                      | 5                    | 2 (40.0)  | 3 (60.0)          | 1.00             |         |
| G3                                      | 14                   | 6 (54.5)  | 5 (45.5)          |                  |         |
| G4                                      | 1                    | 1 (100)   | 0 (0.00)          |                  |         |
| Tumor percentage (%)                    |                      |           |                   |                  |         |
| Mean ± SD                               | 39.3 ± 19.2          | 49.4 ± 21.6| 29.9 ± 12.5       | 0.040            |         |
| Index diameter                          |                      |           |                   |                  |         |
| Mean ± SD                               | 2.00 ± 0.48          | 2.19 ± 0.55| 1.79 ± 0.39       | 0.110            |         |
| Margin status in RP, n (%)              |                      |           |                   |                  |         |
| Negative                                | 4                    | 1 (33.3)  | 2 (66.4)          | 0.576            |         |
| Positive                                | 16                   | 8 (57.2)  | 6 (42.8)          |                  |         |
| Perineural invasion, n (%)              |                      |           |                   |                  |         |
| Intraprostatic                          | 8                    | 3 (50.0)  | 3 (50.0)          | 1.00             |         |
| Intraprostatic + extraprostatic         | 12                   | 6 (54.5)  | 5 (45.5)          |                  |         |
| Lymphovascular invasion, n (%)          |                      |           |                   |                  |         |
| No                                      | 10                   | 2 (25.0)  | 6 (75.0)          | 0.131            |         |
| Yes                                     | 9                    | 6 (75.0)  | 2 (25.0)          |                  |         |
| Unknown                                 | 1                    | 1 (100)   | 0 (0.00)          |                  |         |
| Pathologic stage, pT, n (%)             |                      |           |                   |                  |         |
| 2c-3a                                   | 7                    | 1 (16.7)  | 5 (83.3)          | 0.049            |         |
| 3b                                       | 13                   | 8 (72.7)  | 3 (27.3)          |                  |         |
| Pathologic stage, pN, n (%)             |                      |           |                   |                  |         |
| 0                                        | 4                    | 0 (0.00)  | 4 (100)           | 0.029            |         |
| 1                                        | 16                   | 9 (69.2)  | 4 (30.8)          |                  |         |
| Fusion, n (%)                            |                      |           |                   |                  |         |
| Positive                                | 10                   | 4 (50.0)  | 4 (50.0)          | 1.00             |         |
| Negative                                | 10                   | 5 (55.6)  | 4 (44.4)          |                  |         |
| ERG, n (%)                               |                      |           |                   |                  |         |
| Positive                                | 14                   | 6 (54.6)  | 5 (45.4)          | 1.00             |         |
| Negative                                | 6                    | 3 (50.0)  | 3 (50.0)          |                  |         |
| EZH2, n (%)                              |                      |           |                   |                  |         |
| Positive                                | 15                   | 5 (41.7)  | 7 (58.3)          | 0.294            |         |
| Negative                                | 5                    | 4 (80.0)  | 1 (20.0)          |                  |         |
| NKX3.1, n (%)                            |                      |           |                   |                  |         |
| Positive                                | 9                    | 4 (57.2)  | 3 (42.8)          | 1.00             |         |
| Negative                                | 11                   | 5 (50.0)  | 5 (50.0)          |                  |         |
| SPINK-1, n (%)                           |                      |           |                   |                  |         |
| Positive                                | 12                   | 4 (36.4)  | 7 (63.6)          | 0.131            |         |
TABLE 2 (Continued)

| Characteristics       | Full cohort (n = 20) | BCR time*                  |
|-----------------------|---------------------|----------------------------|
|                       |                     | ≤3 months (n = 9) | >3 months (n = 8) | p-value       |
| Negative              | 8                   | 5 (83.3)          | 1 (16.7)         |              |
| Concordance pattern, n (%) |          |                     |                |
| Concordant            | 3                   | 2 (66.7)          | 1 (33.3)        | 1.00          |
| Partially concordant  | 6                   | 2 (40.0)          | 3 (60.0)        |              |
| Discordant            | 11                  | 5 (55.6)          | 4 (44.4)        |              |

Abbreviations: IQR, interquartile range; RP, radical prostatectomy.
*In 3 cases out of 20, it was not possible to establish BCR.
<0.05.

3.2 Determination of molecular subtypes and molecular patterns by sample and patient

The molecular subtypes analyzed in this study were: ERG(+), EZH2, NKX3.1, and SPINK-1. The ERG(+) subtype includes the presence of the TMPRSS2-ERG fusion and/or highly expressed ERG mRNA. Since no mutations were found in the SPOP gene, this was not considered for analysis. A total of 83 samples (70 prostatic foci and 13 LN) obtained from 20 patients were analyzed. All the results by patient and sample are shown in Table 3.

Figure 2A shows the presence of different molecular subtypes by samples and divided by the presence/absence of the ERG(+) subtype, which is considered the most frequent and most widely studied finding in PCa. The presence of each subtype in the samples analyzed ranges from 41% to 60%; 50 samples (60%) were considered ERG(+), 45 (54%) were EZH2, 38 (46%) were NKX3.1, and 34 (41%) were SPINK-1. A nonmolecular subtype was identified in 7 samples. The presence of each subtype in the samples was predominantly in combination with ERG(+), ranging from 60% to 74%: 27 (60%) of EZH2 samples, 27 (71%) of NKX3.1 samples, and 25 (74%) of SPINK-1 samples. ERG(+) alone was found in only 1 sample (Figure 2A).

To establish patient subtypes, a patient is considered positive for a subtype when they have at least one sample with the subtype present. The presence of each subtype in the samples analyzed ranges from 75% to 85%; there were 17 patients (85%) with ERG(+), 17 (85%) with EZH2, 15 (75%) with NKX3.1, and 15 (75%) with SPINK-1. The presence of each subtype in patients was predominantly in combination with ERG(+), ranging from 82 to 93%; ERG(+) alone was not found in any patient (Figure 2B).

When considering the four subtypes simultaneously, 15 molecular patterns were identified in the samples, with the frequency of these patterns ranging from 1% to 13%; 8 molecular patterns were ERG(+), and they were found in 50 samples (60%), while 7 were ERG(-) in 33 samples (40%) (Figure 2A,C). The simultaneous presence of ERG(+)EZH2/NKX3.1, which are related among them, was found in 11 samples (13%); 8 samples (9.6%) with this pattern and 3 samples (3.6%) also including SPINK-1 ERG(+)/EZH2/NKX3.1/SPINK-1. The most frequent pattern was EZH2 alone (13%), followed by ERG(+)/EZH2/SPINK-1 and ERG(+)/NKX3.1 (11% each).

A total of 8 different molecular patterns were found in patients, with a frequency ranging from 5% to 40%; 5 molecular patterns were ERG(+), and they were found in 17 patients (85%), while 3 patterns were ERG(−) in 3 patients (15%) (Figure 2B,C). The simultaneous presence of ERG(+)/EZH2/NKX3.1, which are related among them, was found in 11 patients (55%): 3 samples (15%) with this pattern and 8 samples (40%) also including SPINK-1 ERG(+)/EZH2/NKX3.1/SPINK-1. The latter was the most frequent pattern in patients. The 3 molecular patterns with ERG(−) are present in only 1 patient each (5%), but all have EZH2 subtype (Figure 2B,C).

3.3 Determination of concordance for each molecular alteration evaluated in prostatic foci

Aiming to establish the clonal origin of the different prostatic foci analyzed in each of the 20 patients, the concordance of molecular characteristics was assessed with respect to the presence or absence of the TMPRSS2-ERG fusion and/or highly expressed ERG mRNA. Since no mutations were found in the SPOP gene, this was not considered for analysis. A total of 8 different molecular patterns were found in patients, with a frequency ranging from 5% to 40%; 5 molecular patterns were ERG(+), and they were found in 17 patients (85%), while 3 patterns were ERG(−) in 3 patients (15%) (Figure 2B,C). The simultaneous presence of ERG(+)/EZH2/NKX3.1, which are related among them, was found in 11 patients (55%): 3 samples (15%) with this pattern and 8 samples (40%) also including SPINK-1 ERG(+)/EZH2/NKX3.1/SPINK-1. The latter was the most frequent pattern in patients. The 3 molecular patterns with ERG(−) are present in only 1 patient each (5%), but all have EZH2 subtype (Figure 2B,C).
|    | Concordant | Partially concordant | Discordant |
|----|------------|----------------------|------------|
|    | Dx | F | Relative expression | Dx | F | Relative expression | Dx | F | Relative expression |
| 1  | GL3 | + | 719.8 | 131.6 | 124.8 | 189.7 | GL3 | + | 7.50 | 3.95 | 7.38 | 0.01 | HGPIN | + | 25.67 | 0.25 | 2.45 | 5.87 |
|    | GL4 | + | 433.5 | 158.5 | 256.5 | 680.9 | GL4 | + | 13.06 | 3.23 | 7.30 | 0.01 | GL3 | + | 8.57 | 2.91 | 1.29 | 2.86 |
|    | PMP | + | - | - | - | - | LN | - | 124.77 | 40.68 | 0.04 | 5.67 | GL4 | + | 25.13 | 0.27 | 2.28 | 6.34 |
| 2  | HGPIN | - | 0.03 | 3.46 | 21.55 | 0.18 | PMP | + | + | + | - | PMP | + | - | - | + | + |
|    | GL3 | - | 0.02 | 6.15 | 13.23 | 0.07 | HGPIN | - | 1.44 | 0.38 | 0.08 | 0.42 | HGPIN | - | 0.56 | 3.43 | 0.00 | 0.47 |
|    | GL3 | - | 0.01 | 2.57 | 6.76 | 0.02 | GL3 | - | 213.70 | 64.89 | 100.90 | 24.48 | GL3 | + | 15.41 | 8.25 | 1.84 | 1.89 |
|    | GL4 | - | 0.01 | 2.49 | 5.57 | 0.03 | GL4 | - | 16.63 | 12.33 | 11.96 | 58.20 | GL4 | + | 1.01 | 11.00 | 1.42 | 0.47 |
|    | PMP | - | + | + | - | - | LN | - | 134.16 | 26.75 | 0.01 | 0.17 | LN | - | 1.40 | 3.25 | 0.20 | 4.32 |
| 3  | HGPIN | - | 0.13 | 0.23 | 3.39 | 0.57 | PMP | + | - | + | + | PMP | + | - | + | + | + |
|    | GL3 | - | 0.03 | 0.19 | 1.90 | 0.52 | HGPIN | - | 9.47 | 0.10 | 0.17 | 0.10 | GL3 | - | 13.38 | 0.01 | 1.16 | 2.22 |
|    | GL4 | - | 0.24 | 0.20 | 2.24 | 0.40 | GL3 | + | 14.56 | 0.02 | 0.15 | 0.02 | GL4 | - | 6.16 | 1.91 | 1.73 | 1.13 |
|    | LN | - | 3.33 | 1.74 | 0.00 | 0.43 | GL4 | + | 1.34 | 0.04 | 0.07 | 0.01 | LN | + | 12.95 | 2.13 | 0.16 | 1.06 |
|    | PMP | + | + | + | - | - | GL5 | + | 675.40 | 5.47 | 0.43 | 1.08 | PMP | + | + | + | + | + |
| 7  | GL4 | - | 0.45 | 0.21 | 0.84 | 52.39 | HGPIN | - | 0.62 | 0.31 | 4.11 | 0.31 | GL3 | - | 0.20 | 0.08 | 0.31 | 0.08 |
|    | GL4 | - | 0.70 | 0.17 | 0.65 | 12.78 | GL4 | + | 3.43 | 0.30 | 2.11 | 0.48 | GL4 | + | 5.78 | 0.50 | 0.28 | 0.50 |
|    | GL5 | - | 0.26 | 0.26 | 0.48 | 11.39 | GL5 | - | 0.63 | 0.09 | 0.15 | 14.64 | LN | + | 1.27 | 0.06 | 0.30 | 0.21 |
|    | GL5 | + | 1.40 | 0.08 | 0.18 | 39.82 | GL5 | + | - | + | + | PMP | + | - | + | - | - |
|    | LN | - | 1.22 | 0.04 | 0.00 | 1.34 | HGPIN | - | 2.16 | 0.81 | 0.22 | 0.64 | HGPIN | - | 6.37 | 1.09 | 0.61 | 7.93 |
|    | PMP | + | - | + | + | + | GL3 | + | 5.78 | 0.50 | 0.28 | 0.50 | GL3 | + | 43.55 | 0.44 | 0.30 | 1.35 |
| 8  | HGPIN | - | 0.35 | 191.10 | 0.40 | 3.40 | GL4 | + | 5.64 | 0.97 | 0.49 | 0.59 | GL4 | + | 5.78 | 0.50 | 0.28 | 0.50 |
|    | GL3 | - | 0.04 | 74.14 | 1.08 | 2.23 | GL4 | + | 37.50 | 2.08 | 0.74 | 0.73 | GL4 | + | 37.50 | 2.08 | 0.74 | 0.73 |
|    | GL3 | - | 0.68 | 37.90 | 2.73 | 0.51 | PMP | + | + | + | + | PMP | + | + | + | + | + |
|    | GL4 | - | 0.29 | 40.07 | 2.21 | 0.07 | PMP | + | + | + | + | PMP | + | + | + | + | + |
|    | GL4 | - | 0.26 | 27.84 | 2.70 | 0.02 | PMP | + | + | + | + | PMP | + | + | + | + | + |
|    | PMP | - | + | + | + | + | PMP | - | + | + | + | PMP | - | + | + | + | + |
| 9  | HGPIN | - | 0.84 | 48.16 | 0.99 | 55.39 | HGPIN | - | 0.37 | 0.57 | 1.38 | 0.97 | HGPIN | - | 0.37 | 0.57 | 1.38 | 0.97 |
|    | GL3 | - | 0.01 | 3.15 | 11.73 | 0.05 | GL3 | + | 13.55 | 1.05 | 0.54 | 2.02 | GL3 | + | 13.55 | 1.05 | 0.54 | 2.02 |
|    | GL3 | - | 0.04 | 2.65 | 2.38 | 0.07 | GL4 | + | 11.69 | 0.62 | 0.17 | 1.74 | GL4 | + | 11.69 | 0.62 | 0.17 | 1.74 |
|    | GL4 | - | 0.14 | 13.43 | 3.84 | 0.58 | GL5 | + | 111.74 | 1.44 | 1.18 | 2.10 | GL5 | + | 111.74 | 1.44 | 1.18 | 2.10 |
|    | PMP | - | + | - | + | + | LN | - | 8.91 | 0.12 | 0.07 | 2.74 | LN | - | 8.91 | 0.12 | 0.07 | 2.74 |
| 17 | HGPIN | - | 1.98 | 2.07 | 0.61 | 0.00 | PMP | + | + | + | + | PMP | + | + | + | + | + |
Among the cases with a concordant pattern, the most common was overexpression of ERG (73%), EZH2 (70%), and NKX3.1 (56%) (pattern A), as well as the absence of fusion (82%) and underexpression of SPINK-1 (75%) (pattern B). The most common pattern for partially concordant cases was pattern C: acquisition of the fusion (100%) and progressive increase in gene expression of ERG (57%), EZH2 (75%), NKX3.1 (63%), and SPINK-1 (67%) (Figure 1F).

### 3.4 Determination of concordance in prostatic foci for all molecular alterations evaluated in each patient

After analyzing concordance for each molecular characteristic separately, concordance was evaluated simultaneously for all molecular characteristics analyzed in this study in prostatic foci (n = 70) (Table 3). A total of 3 out of 20 patients were concordant (15%), 6 patients were partially concordant (30%), and 11 patients were discordant (55%), which shows that heterogeneity and the possibility of multiclonal origin increases from 15 to 30% by molecular characteristic analyzed separately to 55% when the five molecular characteristics are analyzed simultaneously.

Intratumoral concordance was also assessed in foci with the same GL score in 6 patients who had samples with more than one focus with the same GL. Although the foci of the same GL in patients 1, 2, and 9 showed concordance, there were differences in patients 7, 10, and 8 in the presence of fusion and expression of ERG, NKX3.1, and SPINK-1, which indicates intratumoral heterogeneity (Table 3).

### 3.5 Molecular alterations in metastasis concerning prostatic foci

Metastatic material from the lymph nodes of 13 patients was analyzed (Table 3). None of the LN analyzed showed the same pattern.
as any of the other foci examined in the prostate of each patient, which is the expected outcome due to the subclonal events suffered by the cells to metastasize. To establish the possible origin of the metastases found in the LN of the 13 cases, when this material was available, the molecular patterns of the LN and prostatic lesions were compared to identify how the lesion could have caused these metastases to LN.

Table 4 shows the differences found between LN and lesion(s) most similar to LN in each of the 13 cases. Noticeably, in most cases—6 out of 13 (46.2%)—the LN differs from the most similar focus by a single molecular characteristic. The table shows that in 2 of the 6 cases that behave this way, two different foci have the same molecular pattern (cases 7 and 17). In 4 of the 13 cases analyzed (30.7%), two differences were found; in 2 cases (cases 11 and 18), two foci with a different pattern had the same level of similarity with the corresponding LN. As for the 2 cases that presented three differences in this type of comparison (cases 3 and 20), they both showed the same number of differences in all prostate foci examined, which indicates that it is impossible to identify correspondence with a particular focus. A single case presented four differences regarding the PCa foci with different patterns in each (case 19).

In their great majority, the differences found are those associated with progression, increase in the expression of ERG, EZH2, SPINK-1, and decrease of NKX3.1. As for fusion, it is more common to lose it than gain it, which initially would not be expected; however, in the 4 cases with loss of fusion (cases 11, 14, 16, and 18), the ERG expression is not decreased.

### 3.6  Histopathological and clinical significance of the presence of ERG, EZH2, SPINK-1, and NKX3.1 subtypes

Figure 3A shows the frequency of molecular subtypes by histopathological group. All analyzed subtypes were detected starting from the preneoplastic lesion (HGPIN) in a range of 47%–53%, but the SPINK-1 subtype had a low percentage (18%). In LN, the range was between 54% and 92%; the presence of each subtype increased compared to HGPIN, where the most frequent subtypes were ERG and NKX3.1 (92% each). It is clear that ERG overexpression increases in frequency from HGPIN to LN (53%–92%).

Figure 3B shows the frequency of 15 molecular patterns by histopathological group. A noteworthy finding was that pattern heterogeneity decreased with lesion progression; while 7 molecular patterns were found in HGPIN, 9 in GL3, and 11 in GL4, only 5 were found in GL5 and GL6 in LN. All LN and GL5 have a molecular pattern, but in 7 foci—4 HGPIN (24%), 1 GL3 (4%), and 2 GL4 (9%)—no pattern was found, which probably reflects different patterns not identified with the molecular subtypes analyzed here and increased heterogeneity in

![Figure 2](image-url)
these histopathological groups. The molecular pattern with the 4 PCa subtypes (ERG(+) / EZH2 / NKX3.1 / SPINK-1) was detected only in LN, in 23% of them. The constant in all but one LN molecular pattern was the NKX3.1 subtype.

Association between relative gene expression of ERG, EZH2, SPINK-1, and NKX3.1 and histopathological group was also evaluated. The distribution of the relative gene expression of ERG, EZH2, and SPINK-1 showed no significant differences between histopathological groups, but the relative gene expression of NKX3.1 showed a statistically significant difference between groups (p < 0.001). A lower expression in LN samples with respect to all the other groups (LN vs. HGPIN, p = 0.0015; LN vs. GL3, p < 0.001; LN vs. GL4, p < 0.001; LN vs. GL5, p = 0.0236), and between GL5 compared to GL4 samples (p = 0.0071), indicates that there is an inversely proportional association between PCa progression (HGPIN, GL3, GL4 vs. GL5, LN) and the expression levels of NKX3.1 (Figure 4).

An analysis was carried out by comparing clinicopathological characteristics, molecular subtypes, and concordance patterns in all 20 patients by BCR time, categorized as ≤3 and >3 months (Table 2). A statistically significant association was evidenced between BCR time with tumoral percentage (p = 0.040), pT (p = 0.049), and pN (p = 0.029), but no association was found between BCR time and molecular subtypes or molecular concordance. A comparison of the clinicopathological characteristics by molecular subtype with ERG, EZH2, NKX3.1, and SPINK-1 subtypes did not show any statistically significant association (Tables S1–S4).

4 | DISCUSSION

Although PCa is a multifocal disease, it is not clear to what extent molecular heterogeneity in these foci is a determining factor in the natural history of the disease or its role as a prognostic molecular biomarker for this disease. The novelty of this work lies in the infrequency of molecular analyses of many foci in the same PCa with different lesion degrees (HGPIN, GL3, GL4, GL5, and LN). This approach allowed contributing to a better understanding of the molecular heterogeneity of a multifocal disease. In this study, five molecular subtypes associated with PCa were analyzed: ERG(+), EZH2, NKX3.1, SPINK-1, and SPOP mutation, to establish their association with different lesion degrees and regional metastasis in LN in 20 PCa patients with poor prognosis, as well as the clonal origin of the foci in these patients. The ERG(+) subtype includes the TMPRSS2-ERG fusion and ERG overexpression. Since no mutations were identified in the SPOP gene, which may be due to sample size, this was not considered for analysis.

Initially, when determining the presence of molecular subtypes, the ERG(+) subtype was found in 60% of the samples and in 85% of
The previously reported frequency of this subtype for PCa in Caucasian men was between 40% and 66%, similar to what was found in the samples, but our frequency is higher in patients, probably because the present study includes HGPIN and LN, as well as various GL foci that previous studies did not analyze.

The reported EZH2 subtype was identified in 54% of the samples and 85% of the patients, which was more frequently associated with the ERG(+) subtype (60% in samples and 82% in patients). All patients with ERG(−) (18%; n = 3) had the EZH2 subtype. In the literature, the EZH2/ERG-independent subtype has been reported in 30% of the patients. The molecular subtypes and molecular patterns analyzed by histopathological group. (A) Frequency of PCa molecular subtypes. (B) Frequency of molecular patterns by histopathological group. ** The molecular patterns were organized so that they allow showing less heterogeneity in more advanced groups (GL5 and LN).

### Table 1: Molecular Subtypes and Patterns

| Molecular Pattern | HGPIN | GL3 | GL4 | GL5 | LN |
|-------------------|-------|-----|-----|-----|----|
| All ERG(+)        | 9 (53%) | 12 (52%) | 13 (57%) | 4 (57%) | 12 (92%) |
| All ERG(-)        | 8 (47%) | 11 (48%) | 10 (43%) | 3 (43%) | 1 (8%) |
| ERG(+)/NKK3.1/SPINK-1 | 2 (12%) | - | 2 (9%) | 1 (14%) | 4 (31%) |
| ERG(+)/EZH2/NKK3.1 | 4 (24%) | 2 (9%) | 2 (9%) | - | 1 (8%) |
| ERG(+)           | 2 (12%) | 1 (4%) | 3 (13%) | - | 1 (8%) |
| EZH2/NKK3.1      | 1 (6%) | - | - | - | 1 (8%) |
| NKK3.1/SPINK-1   | -      | 5 (22%) | 2 (9%) | 2 (29%) | - |
| NKK3.1           | -      | 3 (13%) | - | 1 (14%) | - |
| EZH2             | 1 (6%) | 5 (22%) | 5 (22%) | - | - |
| ERG(+)           | 1 (6%) | 2 (9%) | 3 (13%) | - | - |
| EZH2/NKK3.1/SPINK-1 | - | 1 (4%) | 1 (4%) | - | - |
| ERG(+)           | -      | - | 1 (4%) | - | - |
| SPINK-1          | -      | - | 1 (4%) | - | - |
| EZH2/SPINK-1     | 2 (12%) | 1 (4%) | - | - | - |
| None             | 4 (24%) | 1 (4%) | 2 (9%) | - | - |
cases by immunohistochemistry (IHC), which is similar to our results (40% in samples, with a decrease to 18% in patients). Nevertheless, it is important to note that our study analyzed many foci per patient, which, given the heterogeneity of PCa, increases the probability of finding an ERG+ focus.

With respect to the NKX3.1 subtype, 46% of the foci and 75% of the patients had this subtype, and the presence of NKX3.1 subtype was more frequent in combination with ERG+ (in 71% of the foci and 93% of the patients). Previous literature has not reported a frequency for this PCa subtype to compare our results. Given the relationship among the expressions of ERG, EZH2, and NKX3.1, these three subtypes (ERG+/EZH2/NKX3.1) were simultaneously present in 11 samples (13%) and in 11 patients (55%). It is important to highlight that these subtypes are also found separately, which indicates that the expression of these genes is more complex than the sole reciprocal relationship among them.

The SPINK-1 subtype was found in 41% of the samples and 75% of the patients. Although most studies reported this subtype in ERG- PCa in approximately 10%-15% of the cases, SPINK-1 was present predominantly in combination with ERG+ in our study (74% of SPINK-1 samples and 87% of the patients). Although SPOP mutation has been linked to SPINK-1 overexpression in various studies, mutations were not found in any foci.

Considering the simultaneous presence/absence of the four subtypes, more molecular patterns (15 patterns) were identified in the samples than in patients (8 patterns). It is noteworthy that the most frequent pattern in samples was an ERG- pattern: EZH2/ERG-independent, which was found in 11 samples (13%); although this number is not too high, there were 3 more patterns that ranged between 10% and 11%. In patients, the most common pattern (40%) was the simultaneous presence of the four subtypes (ERG+/EZH2/NKX3.1/SPINK-1), which is a relevant finding.

Similarly, a comparative analysis was carried out between clinicopathological characteristics, molecular subtypes, and BCR time in the 20 patients included in the study. No association was found between BCR time in patients, categorized as ≤3 and >3 months, and molecular subtypes or molecular concordance. Clinicopathological characteristics and the molecular subtype were compared with ERG, EZH2, NKX3.1, and SPINK-1 subtypes, which did not show any statistically significant association.

It is not surprising that the frequency of the analyzed subtypes increased when analyzed by patient (75%-85%) with respect to foci (41%-60%), while the presence of molecular patterns decreased. This is because most of the literature only analyzes one focus (index tumor) to establish subtypes, while the present study considers that a subtype is present when it is detected at least in one sample of a
patient and includes a range of 3–5 samples (HGPIN, PCa foci, LN) from each patient. Molecular patterns decrease in patients because patterns with the presence of more subtypes increase when summing the presence of the subtypes in all samples analyzed by case.

The presence of molecular alteration of a subtype in at least one sample is significant because a clone with a specific subtype can be the most aggressive one and can determine prognosis and response to therapies. Thus, considering more adaptive tumoral cells, the methodology employed in this study to establish the presence of a subtype in LN is unusual but very significant and increases the presence of subtypes found in patients. The analyzed nonmolecular subtype was identified in 7 prostatic foci (8%); nevertheless, there are other molecular subtypes not evaluated in this work, such as mutations in the IDH1 and FOXA1 genes, described in the TCGA study.3 Interestingly, although the TCGA made a complete molecular analysis of PCa, 26% of the cases did not belong to any of the 7 subtypes described by them.

As for the concordance analysis of prostatic foci per patient, concordance among five molecular characteristics (TMPRSS2-ERG, expression of ERG, EZH2, NKX3.1, and SPINK-1) was used to inquire about the clonal origin of PCa. When each molecular characteristic was evaluated separately, total concordance was the predominant state for most of them. Previous studies have found that 50%–60% of multifocal PCa tumors are concordant regarding the presence or absence of the TMPRSS2-ERG fusion, similar to the 55% found in this study.20,34,39 There are no reports of concordance among foci regarding the expression of ERG, NKX3.1, or EZH2.

The second most frequent pattern for these four alterations was partially concordant cases (20%–45%). Among these, the molecular characteristics vary in relation to disease progression, as observed in subclonal alterations acquired during tumor cell evolution. Of the 5 molecular characteristics, 4 were alterations considered PCa subtypes that appear with disease progression (e.g., increased ERG expression that characterizes the ERG(-) subtype of late appearance). A progressive decrease in gene expression related to PCa was less frequent, only in NKX3.1. These results are not comparable with previous literature since other studies of several foci only analyzed 2 or 3 tumor samples without considering HGPIN and did not present results regarding Gleason score per focus.

Total molecular concordance, which includes concordant and partially concordant cases, was high (70%–90%) for each alteration and supports a monoclonal origin with subclonal events; between 15% and 30% of the cases show discordance, which could be associated with a multiclonal origin. When analyzing all alterations simultaneously per patient, the degree of discordance (55%) was higher than when alterations were evaluated separately (10%–30%), as expected. Other studies have reported a decrease in molecular concordance as the number of analyzed molecular alterations increases; in particular, the whole genome or transcriptome shows that the foci vary widely in the same patient.40–42 However, based on the results of our study, it cannot be discarded that some multifocal PCa may have a monoclonal and others a multiclonal origin; that is, there is no single origin for all PCa.

It is noteworthy that there were intratumoral differences in foci with the same GL in 3 out of 6 patients who had samples with more than one focus with the same GL. Nevertheless, concordance in the other 3 patients supports the probability of a monoclonal origin. Wei et al.23 showed that the correlation between different PCa foci in an individual is low. These findings highlight the importance of evaluating the role of multiclonality, which presents different alterations and subtypes in the same patient, complicating its application in translational oncology since alterations other than the index tumor may be determining metastases.23,42

LN were not considered for concordance analysis, but it should be noted that none of the 13 LN studied was the same as the other foci analyzed, which underlines that malignant cells need other subclonal events to form metastatic foci in LN. This is in accordance with what was said above about subclonal alterations, which must provide additional advantages for such sophisticated processes as metastasis.40

Regarding the association between subtypes and histopathological groups, it was possible to observe all of them in HGPIN lesions, which leads to the conclusion that all the evaluated alterations are early events in the natural history of PCa. It is worth noting that the constant in all but one LN molecular subtype was the ERG(+) and NKX3.1 subtype (92%). The molecular pattern with the four PCa subtypes (ERG(+)/EZH2/NKX3.1/SPINK-1) was detected only in LN (23%). The 13 analyzed LN had more similarity among them than with the prostatic foci of the same patient. Findings in molecular patterns are in accordance with less heterogeneity in more advanced lesions (GL5 and LN). The result of less intermetastatic heterogeneity was expected when mutations are needed in the same genes to confer the same properties to a metastatic microenvironment at a distant site.

Only the expression of the NKX3.1 gene was significantly different among the histopathological groups. Its expression decreased in more advanced lesions. Thus, the lowest levels were found in the LN group, indicating regional metastasis; this finding is consistent with other reports where NKX3.1 has been found down-regulated in many PCa tissues. It has been reported lost or completely eliminated even in the most dysplastic and metastatic PCa tumors.11,44–46

No association was found between BCR time and clinicopathological characteristics, molecular subtypes, or molecular concordance, which may be due to our sample size and highlights the importance of future studies with larger sample sizes that are more representative.

5 | CONCLUSIONS

The results of this study support the notion that, given the molecular heterogeneity of PCa, the molecular evaluation of a single focus does not provide a complete molecular picture of the disease in each patient. Although it is possible to assign molecular subtypes in this way, this type of analysis hinders the association of molecular subtypes with prognosis and the choice of appropriate treatment.
Therefore, it would have a significant clinical application, but it is necessary to have results from several studies that manage to demonstrate the impact of these molecular subtypes on the prognosis and progression of the disease. All the subtypes examined here are of early onset, starting in preneoplastic lesions. NKKX3.1 expression level is the only molecular characteristic analyzed that shows a progression pattern.

AUTHOR CONTRIBUTIONS
All authors have participated in the research that resulted in this work, as well as in the writing, reviewing, and final approval of this manuscript. Yenifer Yamile Segura-Moreno and Martha Lucía Serrano wrote the main manuscript text and prepared the figures and tables.

ACKNOWLEDGMENTS
We thank the Instituto Nacional de Cancerología (Bogotá, Colombia) for its financial support, and the Universidad Nacional de Colombia for academic contribution, as well as Ricardo Sánchez Pedraza and Marcela Núñez for collaboration with statistical analysis.

CONFLICT OF INTEREST
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
All data generated or analysed during this study are included in this published article.

ETHICS STATEMENT
The present research was approved by the Ethics Committee of the Instituto Nacional de Cancerología (Bogotá, Colombia). Since tissues were obtained as part of routine care and not specifically for this study, and data used in this study were previously anonymized, the Ethics Committee of the Instituto Nacional de Cancerología (Bogotá, Colombia) waived the need for informed consent for this study. All methods complied with relevant guidelines and regulations at the national and international levels.

REFERENCES
1. International Agency for Research on Cancer-World Health Organization. Cancer Today. Lyon: IARC; 2020. https://gco.iarc.fr/today/home.
2. D’Amico AV. Risk-based management of prostate cancer. N Engl J Med. 2011;365(2):169-171. doi:10.1056/NEJMe1103829
3. Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. Cell. 2015;163(4):1011-1025. doi:10.1016/j.cell.2015.10.025
4. Tomlins SA, Alshalahaf M, Davicioni E, et al. Characterization of 1577 primary prostate cancers reveals novel biological and clinical-pathologic insights into molecular subtypes. Eur Urol. 2015;68(4):555-567. doi:10.1016/j.euro.2015.04.033
5. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science. 2005;310(5748):644-648. doi:10.1126/science.1117679
6. Kunderfranco P, Mello-Grand M, Cangemi R, et al. ETS transcription factors control transcription of EZH2 and epigenetic silencing of the tumor suppressor gene NKKX3.1 in prostate cancer. PLoS One. 2010;5(5):e10547. doi:10.1371/journal.pone.0010547
7. Thangapazham R, Saenz F, Katta S, et al. Loss of the NKKX3.1 tumor-suppressor promotes the TMPRSS2-ERG fusion gene expression in prostate cancer. BMC Cancer. 2014;14:16. doi:10.1186/1471-2407-14-16
8. Paju A, Hotakainen K, Cao Y, et al. Increased expression of tumor-associated trypsin inhibitor, TATI, in prostate cancer and in androgen-independent 22Rv1 cells. Eur Urol. 2007;52(6):1670-1679. doi:10.1016/j.eururo.2007.01.096
9. Tomlins SA, Rhodes DR, Yu J, et al. The role of SPINK1 in ETS rearrangement-negative prostate cancers. Cancer Cell. 2008;13(6):519-528. doi:10.1016/j.ccr.2008.04.016
10. Duan S, Pagano M. SPOP mutations or ERG rearrangements result in enhanced levels of ERG to promote cell invasion in prostate cancer. Mol Cell. 2015;59(6):883-884. doi:10.1016/j.molecel.2015.09.003
11. Lapointe J, Li C, Giacominì CP, et al. Genomic profiling reveals alternative native genetic pathways of prostate tumorigenesis. Cancer Res. 2007;67(18):8504-8510. doi:10.1158/0008-5472.CAN-07-0673
12. Varambally S, Dhanasekaran SM, Zhou M, et al. The polycomb protein EZH2 is involved in progression of prostate cancer. Nature. 2002;419(6907):624-629. doi:10.1038/nature00175
13. Bowen C, Buendorf L, Voeller HJ, et al. Loss of NKKX3.1 expression in human prostate cancers correlates with tumor progression. Cancer Res. 2000;60(21):6111-6115.
14. Taris M, Irani J, Blanchet P, Multigner L, Cathelineau X, Fromont G. ERG expression in prostate cancer: the prognostic paradox. Prostate. 2014;74(15):1481-1487. doi:10.1002/pros.22863
15. Grupp K, Diebel F, Sirma H, et al. SPINK1 expression is tightly linked to 6q15- and 5q21-deleted ERG-fusion negative prostate cancers but unrelated to PSA recurrence. Prostate. 2013;73(15):1690-1698. doi:10.1002/pros.22707
16. Flavin R, Pettersson A, Hendrickson WK, et al. SPINK1 protein expression and prostate cancer progression. Clin Cancer Res. 2014;20(18):4904-4911. doi:10.1158/1078-0432.CCR-13-1341
17. García-Flores M, Casanova-Salas I, Rubio-Briones J, et al. Clinicopathological significance of the molecular alterations of the SPOP gene in prostate cancer. Eur J Cancer. 2014;50(17):2994-3002. doi:10.1016/j.ejca.2014.08.009
18. Liu D, Takhar M, Alshalahaf M, et al. Impact of the SPOP mutant subtype on the interpretation of clinical parameters in prostate cancer. JCO Precis Oncol. 2018;2018. doi:10.1200/PO.18.00036
19. Shaoq J, Liu D, Ma X, et al. Prognostic value of the SPOP mutant genomic subclass in prostate cancer. Urol Oncol. 2020;38(5):418-422. doi:10.1016/j.juro.2020.02.011
20. Kidd SG, Bogaard M, Carm KT, et al. In situ expression of ERG protein in the context of tumor heterogeneity identifies prostate cancer patients with inferior prognosis. Mol Oncol. 2022;16(15):2810-2822. doi:10.1002/1878-0261.13225
21. Epstein JL, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol. 2016;40(2):244-252. doi:10.1097/PAS.0000000000000530
22. Wei L, Wang J, Lampert E, et al. Intratumoral and intertumoral genomic heterogeneity of multifocal localized prostate cancer impacts molecular classifications and genomic prognosticators. Eur Urol. 2017;71(2):183-192. doi:10.1016/j.euro.2016.07.008
23. Huang CC, Deng FM, Kong MX, Ren Q, Melamed J, Zhou M. Re-evaluating the concept of “dominant/index tumor nodule” in multifocal prostate cancer. Virchows Arch. 2014;464(5):589-594. doi:10.1007/s00428-014-1557-y
24. Wise AM, Stamey TA, McNeal JE, Clayton JL. Morphologic and clinical significance of multifocal prostate cancers in radical prostatectomy specimens. Urology. 2002;60(2):264-269. doi:10.1016/s0090-4295(02)01728-4
25. Cheng L, Song SY, Prettow TG, et al. Evidence of independent origin of multiple tumors from patients with prostate cancer. *J Natl Cancer Inst*. 1998;90(3):233-237. doi:10.1093/jnci/90.3.233

26. Zhang P, Wang D, Zhao Y, et al. Intrinsic BET inhibitor resistance in SPOP-mutated prostate cancer is mediated by BET protein stabilization and AKT-mTORC1 activation. *Nat Med*. 2017;23(9):1055-1062. doi:10.1038/nm.4379

27. Saramäki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, Visakorpi T. TMPRSS2:ERG fusion identifies a subgroup of prostate cancers with a favorable prognosis. *Clin Cancer Res*. 2008;14(11):3395-3400. doi:10.1158/1078-0432.CCR-07-2051

28. Börno ST, Fischer A, Kerick M, et al. Genome-wide DNA methylation events in TMPRSS2-ERG fusion-negative prostate cancers implicate an EZH2-dependent mechanism with miR-26a hypermethylation. *Cancer Discov*. 2012;2(11):1024-1035. doi:10.1158/2159-8290.CD-12-0041

29. Richter E, Masuda K, Cook C, et al. A role for DNA methylation in regulating the growth suppressor PMEPA1 gene in prostate cancer. *Epigenetics*. 2007;2(2):100-109. doi:10.4161/epi.2.2.4611

30. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. * Genome Biol*. 2002;3(7). doi:10.1016/j.gdb.2002.3.7-research0034

31. Pfaffi MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29(9):e45. doi:10.1093/nar/29.9.e45

32. Tang DG. Understanding and targeting prostate cancer cell heterogeneity and plasticity. *Semin Cancer Biol*. 2022;82:68-93. doi:10.1016/j.semcancer.2021.11.001

33. Zhang W, Wang T, Wang Y, et al. Intratumor heterogeneity and clonal evolution revealed in castration-resistant prostate cancer by longitudinal genomic analysis. *Transl Oncol*. 2022;16:101311. doi:10.1016/j.tranon.2021.101311

34. Barry M, Pernier S, Demichelis F, Rubin MA. TMPRSS2-ERG fusion heterogeneity in multifocal prostate cancer: clinical and biologic implications. *Urology*. 2007;70(4):630-633. doi:10.1016/j.urology.2007.08.032

35. Furusato B, Tan SH, Young D, et al. ERG oncoprotein expression in prostate cancer: clonal progression of ERG-positive tumor cells and potential for ERG-based stratification. *Prostate Cancer Prostatic Dis*. 2010;13(3):228-237. doi:10.1038/pcan.2010.23

36. Furusato B, Gao CL, Ravindranath L, et al. Mapping of TMPRSS2-ERG fusions in the context of multi-focal prostate cancer. *Mod Pathol*. 2008;21(2):177-75. doi:10.1038/modpathol.2800981

37. Fontugne J, Davis K, Palanisamy N, et al. Clonal evaluation of prostate cancer foci in biopsies with discontinuous tumor involvement by dual ERG/SPINK1 immunohistochemistry. *Mod Pathol*. 2016;29(2):157-165. doi:10.1038/modpathol.2015.148

38. Melling N, Thomsen E, Tsourilakis MC, et al. Overexpression of enhancer of zeste homolog 2 (EZH2) characterizes an aggressive subset of prostate cancers and predicts patient prognosis independently from pre- and postoperatively assessed clinicopathological parameters. *Carcinogenesis*. 2015;36(11):1333-1340. doi:10.1093/carcin/bgv137

39. Mehra R, Han B, Tomlins SA, et al. Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. *Cancer Res*. 2007;67(17):7991-7995. doi:10.1158/0008-5472.CAN-07-2043

40. Wei Y, Wu J, Gu W, et al. Prognostic value of germline DNA repair gene mutations in De novo metastatic and castration-sensitive prostate cancer. *Oncologist*. 2020;25(7):e1042-e1050. doi:10.1634/theoncologist.2019-0495

41. Cooper CS, Eeles R, Wedge DC, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat Genet*. 2015;47(4):367-372. doi:10.1038/ng.3221

42. Cyrta J, Prandi D, Arora A, et al. Comparative genomics of primary prostate cancer and paired metastases: Insights from 12 molecular case studies. *J Pathol*. 2022;257(3):274-284. doi:10.1002/path.5887

43. Levf M, Zhao S, Acrxona U, et al. Multifocal primary prostate cancer exhibits high degree of genomic heterogeneity. *Eur Urol*. 2019;75(3):498-505. doi:10.1016/j.eururo.2018.08.009

44. Gurel B, Ali TZ, Montgomery EA, et al. NKX3.1 as a marker of prostatic origin in metastatic tumors. *Am J Surg Pathol*. 2010;34(8):1097-1105. doi:10.1097/PAS.0b013e3181e6cbf3

45. Akamatsu S, Takata R, Ashikawa K, et al. A functional variant in NKX3.1 associated with prostate cancer susceptibility down-regulates NKX3.1 expression. *Hum Mol Genet*. 2010;19(21):4265-4272. doi:10.1093/hmg/ddq350

46. Bethel CR, Faith D, Li X, et al. Decreased NKX3.1 protein expression in focal prostatic atrophy, prostatic intraepithelial neoplasia, and adenocarcinoma: association with Gleason score and chromosome 8p deletion. *Cancer Res*. 2006;66(22):10683-10690. doi:10.1158/0008-5472.CAN-06-0963

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.