Whole blood defensin mRNA expression is a predictive biomarker of docetaxel response in castration-resistant prostate cancer

Manish Kohli¹
Charles YF Young²
Donald J Tindall³
Debashis Nandy¹
Kyle M McKenzie³
Graham H Bevan⁴
Krishna Vanaja Donkena⁵

¹Department of Oncology, ²Department of Urology, ³Department of Geriatric Medicine, Mayo Clinic, Rochester, MN, ⁴University of Rochester Medical Center, Rochester, NY, ⁵Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA

Abstract: This study tested the potential of circulating RNA-based signals as predictive biomarkers for docetaxel response in patients with metastatic castration-resistant prostate cancer (CRPC). RNA was analyzed in blood from six CRPC patients by whole-transcriptome sequencing (total RNA-sequencing) before and after docetaxel treatment using the Illumina’s HiSeq platform. Targeted RNA capture and sequencing was performed in an independent cohort of ten patients with CRPC matching the discovery cohort to confirm differential expression of the genes. Response to docetaxel was defined on the basis of prostate-specific antigen levels and imaging criteria. Two-way analysis of variance was used to compare differential gene expression in patients classified as responders versus nonresponders before and after docetaxel treatment. Thirty-four genes with two-fold differentially expressed transcripts in responders versus nonresponders were selected from total RNA-sequencing for further validation. Targeted RNA capture and sequencing showed that 13/34 genes were differentially expressed in responders. Alpha defensin genes DEFA1, DEFA1B, and DEFA3 exhibited significantly higher expression in responder patients compared with nonresponder patients before administration of chemotherapy (fold change >2.5). In addition, post-docetaxel treatment significantly increased transcript levels of these defensin genes in responders (fold change >2.8). Our results reveal that patients with higher defensin RNA transcripts in blood respond well to docetaxel therapy. We suggest that monitoring DEFA1, DEFA1B, and DEFA3 RNA transcripts in blood prior to treatment will be helpful to determine which patients are better candidates to receive docetaxel chemotherapy.

Keywords: docetaxel, RNA-seq, defensin, targeted RNA-seq, prostate cancer

Introduction

Chemotherapy resistance is a major obstacle to improving the survival outcome of metastatic, castration-resistant prostate cancer (CRPC) patients. The development of effective therapy for CRPC has suffered from ambiguity in accurately defining treatment end points. To help overcome this problem, standardized clinical end points for assessing drug efficacy in CRPC patients have been developed. The Prostate Cancer Clinical Trials Working Group 2 currently recommends time-to-event end points for Phase II drug development trials and overall survival for Phase III trials at this stage.¹

The acceptable standard for assessing response to both cytotoxic and noncytotoxic therapies in Phase II trials is now a composite progression-free survival end point in which all assessments included by the composite progression-free survival end point (prostate-specific antigen [PSA] level; bone, soft tissue lesions, and symptom assessments) are performed at the same time point, at the very least at 12 weeks after initiating
treatment for CRPC, thus allowing for adequate drug exposure, while overall survival remains a valid clinical end point for FDA approval at this stage for any new agent.

Docetaxel-based therapy is one of the standard first-line treatments for patients with metastatic CRPC; however, only 50% of men are destined to have a biochemical response after 12-weeks of therapy. There is no known predictive marker of response to docetaxel treatment despite its use as a standard of care for CRPC. Identifying predictive biomarkers would enable us to personalize therapy to select those patients who are more likely to benefit while limiting toxic effects in those patients not deriving benefit from chemotherapy. Although tumor tissue has been the major source for biomarkers at present, however, tumor tissue may at times be insufficient for gene expression analysis after castration. Furthermore, it may not be practical to obtain biopsy tissue from metastatic CRPC patients because of the invasive characteristics for tracking changes during the course of treatment. In addition, initial chemotherapy may itself alter gene expression levels, which may fail to reflect tumor dynamics and changes in drug sensitivity during therapy. No standard approach exists today to predict the response to chemotherapy. The ability to monitor in real time the dynamics of the cancer with noninvasive liquid biopsy biomarkers would facilitate the development of personalized cancer management programs for advanced cancer patients. Recent studies have reported cell-free mRNA of several genes, including CXCR4 and Her2, thymidylate synthase, and BRCA1 mRNA levels in plasma as potential predictive biomarkers for chemotherapy in gastric and non-small-cell lung cancer. Circulatory tumor cell count has been demonstrated to be a predictor of survival and treatment response in CRPC. Considering the shortcomings of PSA as a predictive marker in prostate cancer therapeutics, we examined the feasibility of blood RNA signatures for predicting chemotherapeutic response in CRPC.

With the intent that global changes in gene expression profiles in biofluids may provide value for predicting responses to chemotherapy in metastatic CRPC, in the present study, we applied next-generation RNA-sequencing (RNA-seq) technology to identify blood RNA signatures that distinguish responder (chemosensitive) and nonresponder (chemoresistant) patients of docetaxel therapy. We performed whole-transcriptome sequencing (total RNA-seq) of blood collected before and after receiving docetaxel chemotherapy in responders and nonresponders. Total RNA-seq is a global technique for simultaneously measuring wide dynamic range of the transcriptome, in which a majority of highly expressed genes constitutes the RNA molecules. Total RNA-seq has limited coverage of weakly expressed transcripts, impairing accurate transcript assembly and quantification. Targeted RNA capture and sequencing provides enhanced coverage of the targeted regions and deliver an accurate quantification method for validating expressed transcripts. We further evaluated differentially expressed genes identified from total RNA-seq in an independent cohort of patients by RNA capture-based targeted enrichment sequencing as predictive biomarkers of docetaxel response.

Materials and methods

Patients

Blood specimens from patients with CRPC enrolled between September 2009 and October 2011 to a hospital-based registry on a Mayo Institutional Review Board-approved biospecimen repository were used in this study. Patients with metastatic CRPC received standard-of-care docetaxel chemotherapy. Details of this registry have been reported previously. Blood specimens were collected prospectively before and after four cycles of chemotherapy (each treatment was delivered at 75 mg/m² every 3 weeks). Serum PSA levels were measured both at the time of initiation of docetaxel treatment and after four cycles of chemotherapy. Imaging in this registry was performed every four-treatment cycles as per standard follow-up in all patients receiving chemotherapy together with serial PSA levels every treatment cycle. Patients were categorized into responders and nonresponders by the treating physician who made the decision for changing treatments based on their response assessments performed after four docetaxel treatments. All patients were assessed for PSA at baseline and every 3 weeks thereafter. Bone and soft tissue imaging were performed at baseline and repeated after 12 weeks, thus allowing for adequate drug exposure prior to evaluating response. The Prostate Cancer Clinical Trials Working Group 2 criteria were used for defining PSA response and imaging response.

Patients were deemed to have a PSA progression if PSA increased ≥25% from baseline or nadir and was ≥2 ng/mL after 12 weeks of treatment. Progression based on bone imaging scans was defined as the appearance of greater than or equal to two new lesions. Symptom assessments using quality-of-life scales were not included in the response parameters as they were not considered part of delivering an established standard of care treatment in these patients. Clinical details of 16 patients chosen at random from the registry who had received docetaxel chemotherapy for mCRPC stage (six in the initial discovery cohort and ten in
the follow-up cohort) at two time points of blood collection are presented in Table 1. All patients had bone metastases at the time of treatment initiation and had previously received and progressed on androgen-deprivation therapy, which was continued during chemotherapy.

RNA extraction and whole-transcriptome sequencing

Blood was collected from patients before receiving docetaxel and after four cycles of treatment with PAXgene blood RNA tubes (Qiagen NV, Venlo, the Netherlands). All samples used in this study had RNA integrity numbers of >8.0, which indicates a good integrity of the RNA molecules.\textsuperscript{18} Silica-membrane-based RNA isolation and purification in a spin-column format was performed using PAXgene Blood RNA kit (Qiagen NV).\textsuperscript{19} Whole-transcriptome RNA sequencing (total RNA-seq) was performed in the discovery cohort of six patients. RNA libraries were prepared according to the manufacturer’s instructions for the TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA).\textsuperscript{20} Libraries were loaded onto paired-end flow cells and sequenced as 51×2 paired-end reads on an Illumina HiSeq 2500 with TruSeq SBS version 3 sequencing kits.

Bioinformatic and statistical analysis of the RNA-seq data

After quality control analysis, RNA-seq data were processed by the Mayo Clinic Bioinformatics Core facility. Paired-end reads were aligned by TopHat 2.0.6 against the hg19 genome build using the bowtie 1 aligner option.\textsuperscript{21} To compare the differential RNA-sequence analyses in responders versus nonresponders, GeneSifter software edition version 4.0 (PerkinElmer Inc., Waltham, MA, USA) was used. Data were log-transformed, and the Benjamini and Hochberg correction was performed to determine the falsecovery rate. The likelihood ratio test was used to select differentially expressed genes. Two-way analysis of variance was performed to compare differential gene expression in responders versus nonresponders before and after treatment. Normalization of samples was performed using reads per kilobase per million mapped reads (RPKM).

RNA capture-based targeted enrichment sequencing

Differentially regulated genes from the RNA-seq experiment in the discover cohort were validated by SureSelect custom RNA capture-based targeted enrichment sequencing (Agilent Technologies, Santa Clara, CA, USA) in the independent follow-up cohort of ten patients (before and after four docetaxel treatments). RNA libraries were prepared using NEBNext Ultra RNA Library Prep Kit (New England Biolabs Inc., Ipswich, MA, USA). The total design of the custom capture library covered 178,500 kb of target regions of genes, tiled at two times across the targets. We used 3,162 total baits for 34 targets. Biotinylated RNA library baits were prepared, amplified, hybridized to RNA, and selected using streptavidin-coated beads following manufacturer instructions (Agilent Technologies). Sequencing was performed on an Illumina HiSeq 2000 for paired-end reads. Analysis of the sequencing data was performed following a protocol similar to that described earlier for total RNA-seq.

Results

Demographic characteristics and chemotherapy outcomes of patients for both discovery and follow-up cohorts are described in Table 1. In the discovery cohort, the median number of treatment cycles for responders was 16 and for nonresponders was 6, while the median time to clinical progression by imaging criteria was 14 months for responders and 5 months for nonresponders. In the follow-up cohort, the median number of treatment cycles for responders was 10 and for nonresponders was 4, while the median time to clinical progression by imaging was 12.8 months for responders and 3.1 months for nonresponders.

Identification of differentially expressed genes in the responder and nonresponder patients by total RNA-seq

Total RNA-seq in the discovery cohort of six patients pre- and postdocetaxel therapy revealed that 200 genes were differentially expressed in the responders versus nonresponders before treatment and 247 genes were altered in expression following treatment; these genes were identified using a threshold of 1.5 and a \(P\)-value <0.05. We selected the top 34 genes that were differentially expressed greater than or equal to two-fold in patients responding to docetaxel after four treatments (Figure 1A) for further evaluation.

Validation of genes by targeted RNA capture and sequencing confirms differential expression levels in the responders

Targeted RNA capture sequencing was performed to validate the expression of genes in an independent follow-up cohort of 10 patients. Our validation confirmed two-fold higher expression levels of the transcripts in the responders.
Table 1 Demographic and treatment characteristics of patients with metastatic castration-resistant prostate cancer, used for total RNA sequencing and targeted RNA capture and sequencing

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age at initial diagnosis</th>
<th>Body mass index (kg/m²) at the time of initiation of chemotherapy</th>
<th>Gleason score at initial diagnosis</th>
<th>Initial tumor stage</th>
<th>Time from initial cancer diagnosis to starting ADT (in months)</th>
<th>Time on ADT prior to initiating chemotherapy (in months)</th>
<th>Pre-chemotherapy PSA (ng/mL)</th>
<th>Post chemotherapy PSA after 4 treatment cycles (ng/mL)</th>
<th>Total docetaxel cycles</th>
<th>Clinical progression by imaging (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Responders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>64</td>
<td>41.47</td>
<td>9</td>
<td>T4</td>
<td>1.03</td>
<td>15.67</td>
<td>4.80</td>
<td>0.42</td>
<td>16</td>
<td>15.00</td>
</tr>
<tr>
<td>2</td>
<td>70</td>
<td>31.63</td>
<td>7</td>
<td>T3b</td>
<td>25.10</td>
<td>179.03</td>
<td>47.10</td>
<td>45.60</td>
<td>16</td>
<td>14.03</td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>27.77</td>
<td>5</td>
<td>T3b</td>
<td>106.87</td>
<td>83.10</td>
<td>278</td>
<td>87.20</td>
<td>8</td>
<td>13.43</td>
</tr>
<tr>
<td><strong>Nonresponders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>27.78</td>
<td>8</td>
<td>T3b</td>
<td>62.10</td>
<td>36.63</td>
<td>3.40</td>
<td>4.60</td>
<td>6</td>
<td>3.50</td>
</tr>
<tr>
<td>5</td>
<td>72</td>
<td>31.89</td>
<td>9</td>
<td>T3b</td>
<td>1.20</td>
<td>7.77</td>
<td>0.48</td>
<td>0.10</td>
<td>6</td>
<td>5.00</td>
</tr>
<tr>
<td>6</td>
<td>76</td>
<td>38.20</td>
<td>7</td>
<td>T2a</td>
<td>2.03</td>
<td>92.63</td>
<td>15.50</td>
<td>19.70</td>
<td>8</td>
<td>6.33</td>
</tr>
<tr>
<td><strong>Targeted RNA capture and sequencing in follow-up cohort (n=10)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Responders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>70</td>
<td>27.47</td>
<td>9</td>
<td>T3a</td>
<td>80.93</td>
<td>103.87</td>
<td>24.30</td>
<td>22.10</td>
<td>10</td>
<td>13.83</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>32.11</td>
<td>9</td>
<td>T2c</td>
<td>8.10</td>
<td>57.40</td>
<td>144.00</td>
<td>52.00</td>
<td>10</td>
<td>12.87</td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>23.09</td>
<td>6</td>
<td>T2a</td>
<td>31.83</td>
<td>9.83</td>
<td>3.30</td>
<td>1.00</td>
<td>10</td>
<td>8.00</td>
</tr>
<tr>
<td>4</td>
<td>75</td>
<td>35.28</td>
<td>7</td>
<td>T2a</td>
<td>93.37</td>
<td>9.30</td>
<td>32.40</td>
<td>15.10</td>
<td>11</td>
<td>13.13</td>
</tr>
<tr>
<td>5</td>
<td>61</td>
<td>27.59</td>
<td>9</td>
<td>T3a</td>
<td>6.77</td>
<td>13.73</td>
<td>4.60</td>
<td>1.10</td>
<td>10</td>
<td>12.07</td>
</tr>
<tr>
<td><strong>Nonresponders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>75</td>
<td>31.15</td>
<td>7</td>
<td>T2c</td>
<td>109.10</td>
<td>75.20</td>
<td>84.00</td>
<td>156.00</td>
<td>6</td>
<td>3.50</td>
</tr>
<tr>
<td>7</td>
<td>87</td>
<td>24.77</td>
<td>5</td>
<td>T2a</td>
<td>168.53</td>
<td>177.80</td>
<td>57.60</td>
<td>63.60</td>
<td>4</td>
<td>3.13</td>
</tr>
<tr>
<td>8</td>
<td>67</td>
<td>25.76</td>
<td>9</td>
<td>T2c</td>
<td>129.43</td>
<td>39.30</td>
<td>1.60</td>
<td>2.80</td>
<td>4</td>
<td>3.93</td>
</tr>
<tr>
<td>9</td>
<td>75</td>
<td>26.84</td>
<td>8</td>
<td>T3b</td>
<td>52.13</td>
<td>9.70</td>
<td>12.90</td>
<td>17.90</td>
<td>4</td>
<td>2.83</td>
</tr>
<tr>
<td>10</td>
<td>73</td>
<td>32.41</td>
<td>8</td>
<td>T2c</td>
<td>5.03</td>
<td>202.80</td>
<td>14.40</td>
<td>17.30</td>
<td>4</td>
<td>2.90</td>
</tr>
</tbody>
</table>

**Abbreviations:** PSA, prostate-specific antigen; ADT, androgen deprivation therapy.
versus nonresponders for 13 of the 34 genes from the total RNA-seq (Figure 1). The log$_2$-transformed RPKM levels of the 34 genes by total RNA-seq and 13 genes by targeted RNA capture and sequencing are depicted in the clustering diagram (Figure 1A and B).

Targeted RNA capture and sequencing in the follow-up cohort showed 2.5 fold increase in the pre-treatment expression of $\alpha$-defensin genes DEFA1, DEFA1B, and DEFA3 genes in responders compared with nonresponders (Table 2). Two-way analysis of variance analysis showed a significant P-value of 0.025, 0.031, and 0.033 for DEFA1, DEFA1B, and DEFA3 RNA transcripts, respectively, in the responder and nonresponder groups. In addition, docetaxel chemotherapy increased transcripts levels in the responders for these genes 2.8 fold after four treatments compared to nonresponding patients. Reads covering the exon junctions for DEFA1, DEFA1B, and DEFA3 genes that were identified in the total RNA-seq were also found with targeted RNA capture sequencing. In addition, there were reads mapped to novel exon junctions for DEFA1B and DEFA3 genes (Table 3).

Validation of the results further showed two-fold higher expression levels of activating transcription factor 3 (ATF3), cathepsin G (CTSG), GRB2-binding adaptor protein transmembrane (GAPT), keratin 73 (KRT73) and olfactomedin 4 gene (OLFM4) in responders, whereas chemokine receptor 7 (CCR7), kallikrein KLK2, kallikrein KLK3, solute carrier family 12 (SLC12A1), and teashirt zinc finger homeobox 2 (TSHZ2) were two-fold higher in the nonresponders before treatment.

**Discussion**

Gleason score and tumor stage at initial diagnosis and PSA levels at pre- and posttreatments were not helpful to distinguish responders from nonresponders (Table 1). Our observation of higher pretreatment transcript levels of defensins, which increased further after four treatment cycles of docetaxel in responsive patients (Figure 1 and Table 2), suggests a potential tumor-treatment interaction not previously reported. With a much greater number of reads mapped to the exons in the targeted RNA capture sequencing, the twofold difference in the expression levels with a significant P-value was only observed for 13 genes in the follow-up cohort compared to 34 genes in the discovery cohort (Figure 1). Although this pilot study of total RNA-seq has a limited size of study subjects, targeted RNA capture sequencing
confirmed consistent increase in the levels of defensin genes in the responders versus nonresponders.

The high coverage of baits for the selected genes in our validation facilitated better quantitation of reads compared with total RNA-seq. As evident from our results, the reads covering exon junctions for **DEFA3** in the targeted RNA capture sequencing were about ten times higher compared to the total RNA-seq (Table 3), which confirmed that the mapped reads were exclusively RNA transcripts. Defensins are antimicrobial peptides involved in host defense. α-defensins have a distinct role in facilitating differential oncolysis of tumor cells, with **DEFA1** and **DEFA3** directly cytotoxic to tumor cells. **DEFA1** and **DEFA3** encode human neutrophil peptides HNP1, HNP2, and HNP3. In addition, intratumoral expression of HNP1 mediates tumor microvessel normalization and therefore potentiates the therapeutic effect of doxorubin in breast cancer. Organized structured vasculature is a key for efficient delivery of drugs to the tumor. Our observation of a strong association of elevated prechemotherapy α-defensins **DEFA1**, **DEFA1B**, and **DEFA3** expression in patients responding to docetaxel is interesting and suggests further investigation of the role of α-defensin in potentiation of docetaxel effect in CRPC.

In addition to the defensin RNA transcripts, our validation study revealed twofold higher expression levels of **ATF3**, **CTSG**, **GAPT**, **KRT73**, and **OLFM4** in the responders before treatment (Figure 1B). **ATF3** is a tumor suppressor for the major subset of prostate cancers harboring dysfunctional Pten. **CTSG** is an endoprotease that plays an important role in tumor metastasis. **GAPT** participates in hematopoiesis and regulation of immune responses in cancer and inflammation. **KRT73** (also known as K6IRS3) is a class II keratin protein responsible for the structural integrity of epithelial cells. **OLFM4** RNA expression is reduced in prostate cancer tissues. The twofold higher expression levels of these genes in the responder patients compared to nonresponders suggests roles for these genes in response to chemotherapy.

Nonresponder patients showed twofold higher expression levels of **CCR7**, **KLK2**, **KLK3**, **SLC12A1**, and **TSHZ2** before treatment. **CCR7** promotes cell proliferation in several human cancers. Positive staining of **CCR7** has been reported in 88.3% of radical prostatectomy cases. The kallikreins **KLK2** and **KLK3** are a subgroup of serine proteases implicated in carcinogenesis and have potential roles...
as novel cancer biomarkers.\textsuperscript{35} PSA protein is the product of the KLK3 gene used in the diagnosis and monitoring of prostate cancer.\textsuperscript{36} SLC12A1 plays a vital role in the regulation of ionic balance and cell volume.\textsuperscript{37} TSHZ2 is differentially methylated in prostate cancer.\textsuperscript{38} Higher expression of these genes in nonresponder patients indicates that they could be associated with resistance to treatment.

The use of a “liquid biopsy” to assess genetic marker changes after treatment is an attractive option particularly in advanced stage cancer to track disease progression without the need for more invasive and costlier tissue biopsies.\textsuperscript{11} Our findings suggest a promising role for blood defensin levels as predictive biomarkers in CRPC and warrant further expansion of these studies to larger cohorts to determine the significance of defensin levels in docetaxel chemosensitivity. This study however differs from other RNA marker studies whose samples were derived from different biospecimens (eg, RNA extracted from circulating tumor cells),\textsuperscript{39} which demonstrated AR-V7 expression to be associated with resistance to enzalutamide and abiraterone acetate. Nonetheless, both approaches have excellent potential to better inform treatment efficacy. For example, treatments based on blood biomarker levels in CRPC patients with high blood defensins may benefit from docetaxel chemotherapy, while those who show an increase in defensin levels with treatment could benefit from continuation of treatment. On the other hand, lower defensin levels either initially or during treatment may be indicative for switching treatment earlier and avoiding chemotherapy toxicity. In summary, our results reveal that patients with higher defensin gene expressions in blood respond well to docetaxel chemotherapy. We suggest that α-defensin levels in CPRA patients could be used as predictive biomarkers for selection of personalized chemotherapy. In this pilot study, we have discovered a number of promising biomarkers. Our goal is to validate these preliminary findings in a larger cohort of patients in a prospective trial that can evaluate the heterogeneity of CRPC.

**Acknowledgments**

The authors thank the Mayo Clinic Medical Genome Core Facility for their support and help in performing the sequencing. This work is supported in part by the John P Vaile Development Fund and Joseph Gassner Development Fund.

**Disclosure**

Krishna Vanaja Donkena is supported by the grant from the American Cancer Society RSG-09-175-01-CCE. Donald J Tindall is supported by a grant from the T.J. Martell Foundation and Manish Kohli is supported by the Roger Thurn World Health Resources group fund. All authors are cognizant of their inclusion in this paper and agree to the content. There are no other conflicts of interest inherent in this paper.

**References**


