

Targeted sequencing of ampliconic gene transcripts from total human male testis RNA

Marta Tomaszekiewicz (✉ mat19@psu.edu)

Makova Lab

Kateryna Makova (✉ kdm16@psu.edu)

Makova Lab

Method Article

Keywords: Ampliconic genes, Y chromosome, transcripts, PacBio sequencing, Illumina sequencing

Posted Date: September 20th, 2018

DOI: <https://doi.org/10.1038/protex.2018.109>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The protocol summarizes all the steps towards the construction of PacBio and Illumina sequencing libraries of transcripts from nine Y chromosome ampliconic gene families starting from total human male testis RNA from two human males. This method is based on PacBio's Isoform Sequencing method \ (Iso-Seq) using the Clontech SMARTer PCR cDNA Synthesis Kit and No Size Selection with some modifications.

Introduction

The nine ampliconic multi-copy gene families \ (including both exons and introns) occupy ~20% of the human Y chromosome ¹. They are located within palindromes or inverted repeats with the exception of TSPY, which makes the sequence identity of their copies very high \ (>99.9%). Such high identity poses a major challenge for sequencing and assembling of ampliconic gene copies, therefore novel methods are being developed for accurate capture of all the sequences within the same gene family ². Studying ampliconic genes at the transcription level is even more important as we can predict their protein-coding capacity and decipher their isoform diversity. To the best of our knowledge, the isoform diversity of the Y chromosome ampliconic gene transcripts in humans has not been previously studied. Hereby, we present an experimental method for targeted sequencing of ampliconic gene transcripts from total human male testis RNA using short \ (Illumina) and long-read \ (PacBio) technologies. While the method was developed for the human Y chromosome ampliconic gene families, it can be applied to other multi-gene families in human and other species. This protocol is designed for processing two samples, however it can be scaled up to any number of samples \ (limited by the number of sample-specific barcodes to use).

Reagents

Step 1: 5 M Ammonium Acetate Glycogen Ethanol -20°C Freezer Step 2: Poly\ (A) Purist MAG Kit Ethanol
Step 3: Clontech SMARTer PCR cDNA synthesis kit Step 4: Clontech SMARTer PCR cDNA synthesis kit
Kapa HiFi PCR Kit 1 kb marker Electrophoresis buffer DNA gel stain Qubit dsDNA BR Assay Kit Agilent
DNA 12000 Kit Step 5: Gene-family-specific primers for amplifying transcripts from nine ampliconic gene
families in human incorporating PacBio sample-specific barcodes \ (See Table 1). Step 6: KAPA HiFi PCR
Kit Deionized water 1kb marker Electrophoresis buffer DNA gel stain Step 7: KAPA HiFi PCR Kit Deionized
water 1kb marker Electrophoresis buffer DNA gel stain AMPure XP beads Step 8: SMRTbell Template
Prep Kit 1.0 Agilent DNA 12000 Kit Step 9: MiSeq Reagent Nano Kit Nextera XT Sample Preparation Kit

Equipment

Step 1: Vortex mixer Centrifuge 1.5 ml DNA LoBind tubes Step 2: Magnetic stand PCR machine or heating
block for heat incubation 1.5 mL DNA LoBind tubes Step 3: PCR machine or heating block for incubation
Step 4: PCR machine Electrophoresis equipment Qubit BioAnalyzer Step 6: PCR machine Electrophoresis
equipment Step 7: PCR machine Electrophoresis equipment Magnetic stand Step 8: Qubit BioAnalyzer

Procedure

Step 1: Precipitation of total RNA (residual salt removal) 1. Add 0.1 volume of 5 M Ammonium Acetate to 5-50 ug of testis RNA. 2. Add 1 µL Glycogen. 3. Add 2.5 volumes of 100% ethanol. 4. Mix thoroughly by vortexing. 5. Precipitate at -20°C overnight. 6. Recover the RNA by centrifugation at 12,000 x g for 30 min at 4° C. 7. Carefully remove and discard the supernatant. 8. Centrifuge the tube briefly for the second time and aspirate away any additional fluid. 9. Add 1 mL 70% ethanol, and vortex the tube a few times. Repellet the RNA by microcentrifuging, for 10 min at 4°C. Remove supernatant carefully not to disturb the RNA pellet. Step 2: Isolation of poly(A) RNA using Poly(A) Purist MAG (Catalog Number: AM1922, Publication Number: 1922, Revision D) Step 3: Generation of first-strand cDNA using Clontech SMARTer PCR cDNA synthesis kit (Catalog Number: 634926, Protocol Number: PT4097-1, Version number: 040114) Step 4: Generation of double-strand cDNA using Clontech SMARTer PCR cDNA synthesis kit (Catalog Number: 634926, Protocol Number: PT4097-1, Version number: 040114) 1. PCR Cycle Optimization for Large-Scale PCR using KAPA HiFi Enzyme 2. Generation of large-scale double-stranded DNA using the optimized cycle number. Step 5: RT-PCR primer design for capturing nine ampliconic gene transcript families from human males For each of the samples, a unique PacBio barcode is added to the forward and the reverse primers according to Guidelines for Using PacBio Barcodes for SMRT Sequencing (<http://www.2einteractive.com/pacbio/Shared-Protocol-PacBio-Barcodes-for-SMRT-Sequencing.pdf>). This is important for the differentiation of transcripts coming from different samples. Step 6: Gradient PCR on human testis cDNA to determine the optimal annealing temperature for each RT-PCR primer pair specific to each multigene transcript family Nine ampliconic gene transcript family targets were amplified using a range of annealing temperatures via Gradient PCR (55-68°C). Subsequently, RT-PCR products were separated with electrophoresis on a 0.8% agarose gel and visualized under the UV light. Distinct bands on the gel allowed us to determine the optimal annealing temperature for capturing each of the multigene transcript family targets (See Table 2). Step 7: Final PCR amplifications of each multigene transcript family targets using a specific annealing temperature (See Table 2), PCR protocol (See Table 3) and RT-PCR product purification using AMPure XP beads Step 8: Preparation of RT-PCR products for PacBio sequencing (RSII) Multigene family RT-PCR products were processed in two groups. Equimolar amounts of RT-PCR products specific to each multigene family and of size less than 1 kb are pooled to the total amount of 2 ug, from which one PacBio SMRTbell library is prepared and sequenced using one SMRT cell using PacBio's Procedure & Checklist – Isoform Sequencing (Iso-Seq) using the Clontech SMARTer PCR cDNA Synthesis Kit and No Size Selection (<https://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-Isoform-Sequencing-Iso-Seq-using-the-Clontech-SMARTer-PCR-cDNA-Synthesis-Kit-and-No-Size-Selection.pdf>). Equimolar amounts of RT-PCR products specific to each gene family and of size greater than 1 kb and smaller than 2 kb have been pooled to the total amount of 2 ug, from which one PacBio SMRTbell library has been prepared and sequenced using one SMRT cell. Step 9: Preparation of RT-PCR products for Illumina sequencing Nextera

XT libraries were constructed from the same RT-PCR products using Nextera XT DNA Sample Preparation Guide (Part# 15031942, Revision C). Equal volumes of bead-normalized libraries are pooled and sequenced on a MiSeq instrument using one MiSeq Reagent Nano Kit, v2 (250x250 paired-end sequencing).

Timing

Step 1 & 2: 1 day Step 3 & 4: 2 days Step 5: 1 week Step 6: 2 days Step 7: 1 day Step 8: 3 days Step 9: 2 days

Troubleshooting

See Table 4.

Anticipated Results

Step 1: Residual salt, if present, is removed from total RNA. Step 2: 1-3% of poly(A) RNA is isolated from total RNA. Step 4: Double-stranded cDNA is generated and appears as a smear from ~0.5 kb to 5 kb of size on 0.8% agarose gel (an example shown in Figure 1). Step 6: A distinct band(s) on an agarose gel indicates the right specific annealing temperature for PCR amplification with a specific RT-PCR primer pair (an example shown in Figure 2). Step 8: 1-2 Gb of data should be generated per one SMRT cell. Step 9: BioAnalyzer profiles of Nextera XT libraries should show a wide distribution of DNA fragment sizes of less than 1 kb (an example shown in Figure 3).

References

1. Skaletsky, H. et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423, 825–837 (2003). 2. Tomaszewicz, M., Medvedev, P. & Makova, K. D. Y and W Chromosome Assemblies: Approaches and Discoveries. *Trends Genet.* 33, 266–282 (2017).

Figures

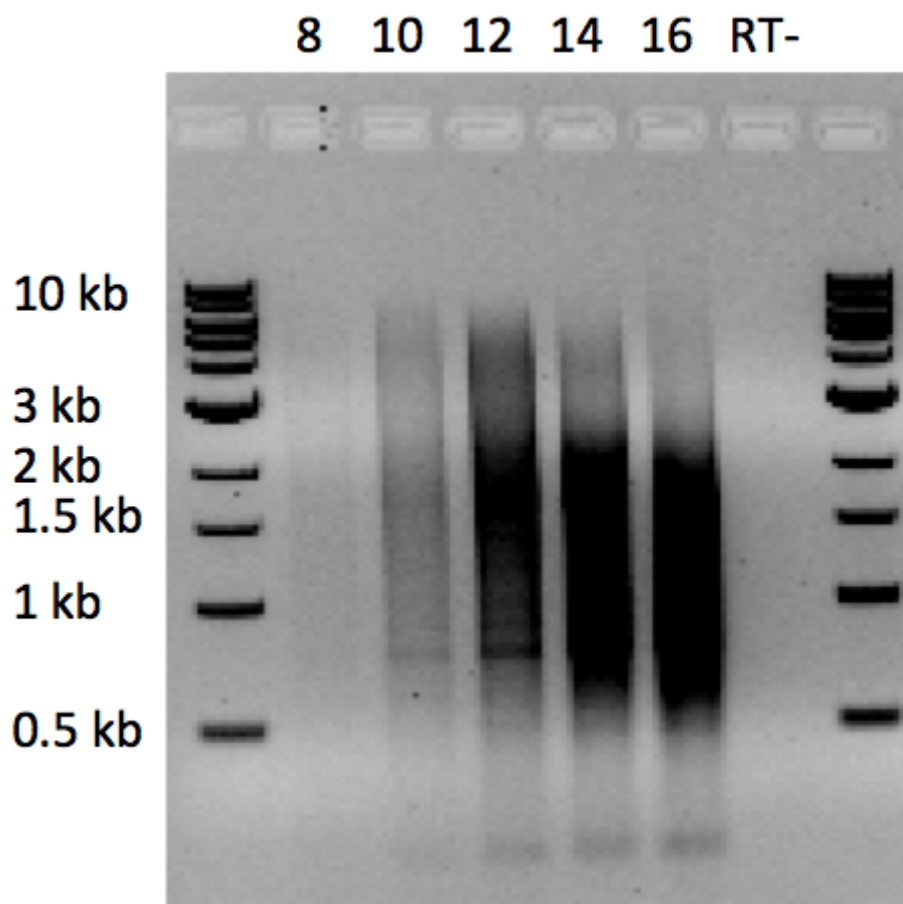


Figure 1

The result of Step 4. Gel electrophoresis of double-stranded cDNA generated using 8-16 PCR cycles (indicated above the lanes) on 0.8% agarose gel using 1 kb NEB marker. PCR with 12 cycles was determined to be optimal for large-scale amplification.

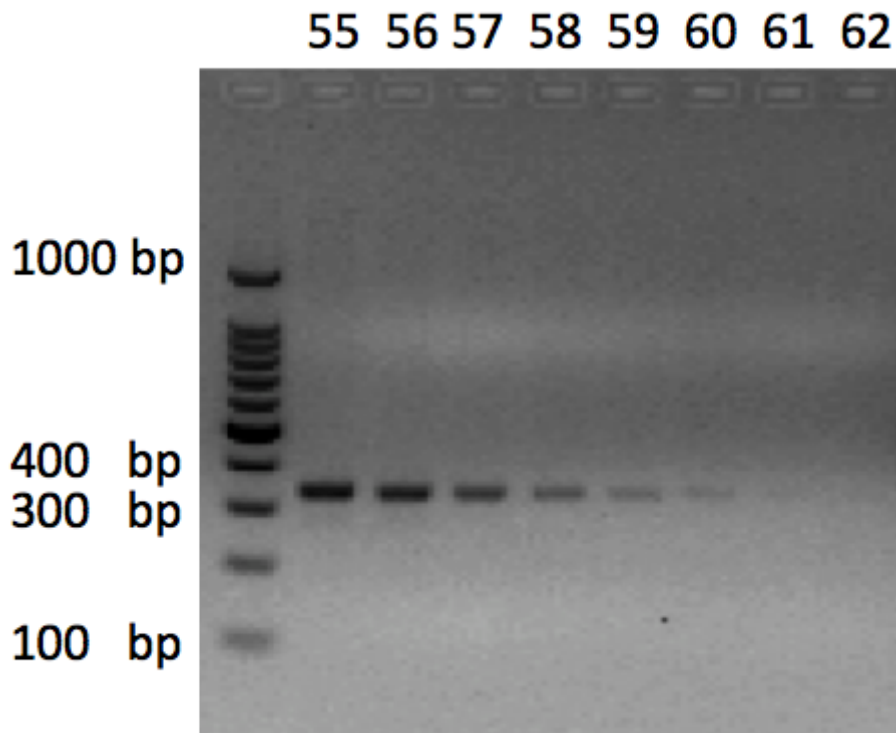


Figure 2

The result of Step 6. Gel electrophoresis of XKRY RT-PCR product amplified using a range of annealing temperatures via Gradient PCR (55-62°C, indicated above the lanes). 55°C was determined to be optimal annealing temperature for capturing XKRY transcripts.

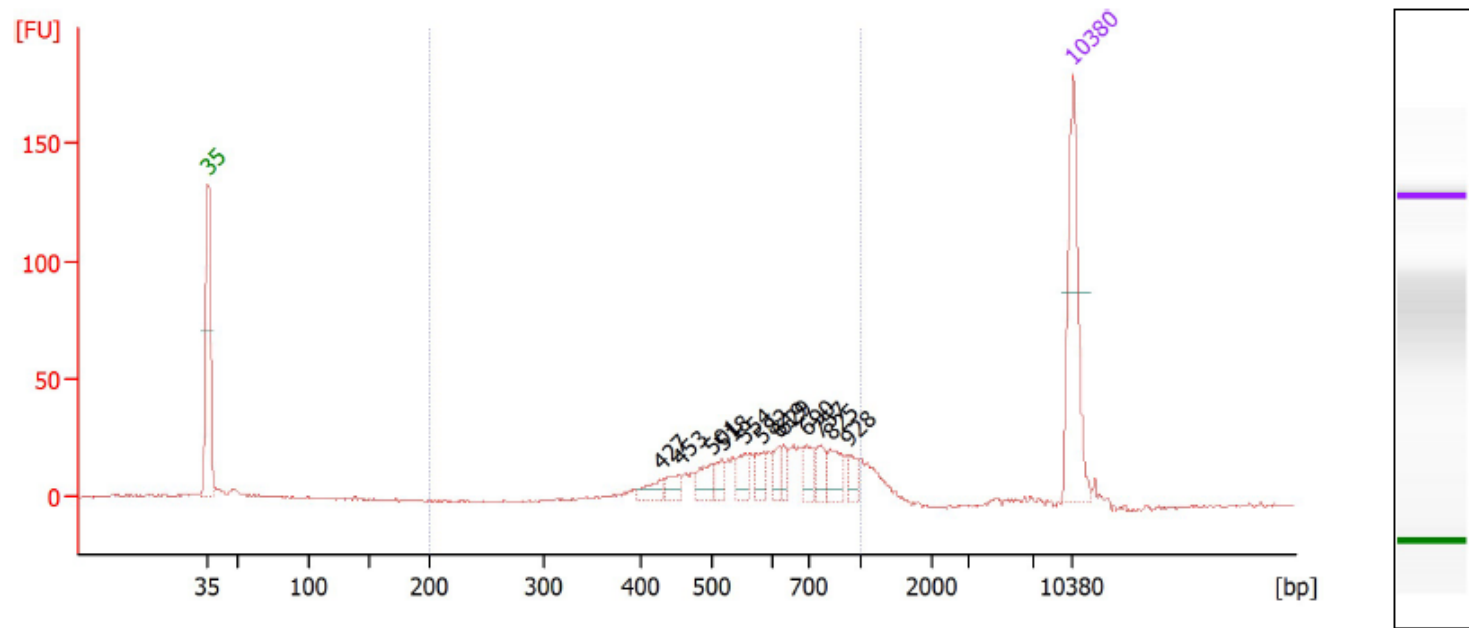


Figure 3

The result of Step 9. Example of Nextera XT library constructed from 1 ng of the CDY1 RT-PCR product.

Supplementary Files

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

- [supplement0.docx](#)
- [supplement0.docx](#)
- [supplement0.docx](#)
- [supplement0.docx](#)