Modification of the Transplex™ WTA2 Amplification Product for Next Generation Sequencing

Ken Heuermann, Brian Ward, Dom Fenoglio

Life Science and Technology Center, 2909 Laclede Avenue, St. Louis, MO 63103

Abstract

Transplex Whole Transcriptome Amplification (WTA2) exponentially amplifies RNA producing a double-stranded cDNA library while precisely maintaining differential levels of individual transcripts in test and reference samples. Though originally designed to amplify nanogram quantities of RNA, Transplex WTA2 has been shown to be exceedingly effective for amplification from damaged RNA template (FFPE and laser captured tissue samples) and single-cell input quantities (picograms). The efficacy of Transplex WTA2 amplification for downstream applications, primarily qPCR and expression microarray analysis, is well-documented. It follows that the utilization of next-generation sequencing for gene expression research and diagnostics would be well served by Transplex amplification of RNA isolated from samples of severely restricted quantity or quality.

Strategies for the integration of Transplex WTA2 with next-generation sequencing are examined, with particular emphasis on elimination of the characteristic fixed primer sequence associated with each amplicon in the amplification library. Removal of these sites will allow direct entry of the resulting product into the sequencing workflow. Methods under consideration will enable the WTA2 amplicon to feed into the current sample prep protocols for the Illumina GA and GAII, SoLiD 5500/5500xl, and Roche-454 GS FLX/Junior platforms.

Introduction

Platform Strengths

Transplex Complete WTA2 kit has been demonstrated to effectively amplify damaged and/or low quantities of RNA template[1-13]. Exponential PCR-based amplification provides both sensitivity and reproducibility, while maintaining relative levels of transcript abundance[1,2,4,8,10-12].

- The Transplex WTA2 library synthesis primer design was improved to provide increased coverage of the transcriptome (over that of WTA1), allowing for enhanced ability to amplify fragmented RNA[14].
- The 3′ component of library synthesis primer, comprised of quasi-random sequence (**green, Transplex WTA2 Workflow**), primes throughout the entire length of its respective template during first- and second strand cDNA synthesis, eliminating the 3′ bias[15,16] characteristic of the "Eberwine" amplification methods[17]. Hence, Transplex WTA2 amplicon is perfectly suited for exon and alternative splicing studies.
- The single universal primer (**blue, Transplex WTA2 Workflow**) is non-self complementary, eliminating potential intra- and intermolecular hybridization of the primer itself[18]. This, combined with amplification conditions and reaction parameters that help to prevent intra-molecular hybridization of the complementary terminal ends of each amplicon during amplification[19], provides for substantially better amplification efficiency over platforms the rely on random "dN" sequence for library synthesis.

Misconceptions

The most common misconception, detracting from the general acceptance of PCR-based amplification strategies, is the comparison of the disjointed increase in relative levels of two different transcripts during amplification in the *same reaction*—due primarily to differences in amplicon length and complexity. *This is not a legitimate argument.* TransPlex WTA2 amplification technology addresses this misconception through the following points:

- The imperative for RNA amplification is the maintenance of relative levels of the same amplicon in, minimally, two samples: the test and reference sample. All amplification methods possess systematic bias, and therefore differential expression, not individual reads, is the true standard. In this way, Transplex WTA is highly reproducible[1,2,4].
- Considering the expression differential between the test and reference sample, PCR amplification is more sensitive, able, to detect stochastic basal expression (eliminated from consideration by statistical analyses)[4].

Applications

- Recent publications have demonstrated the proficiency of the Transplex WTA2 kit for microarray application, including rigorous evaluations of the amplification method, in the case of Gonzalez-Roca et al[4], and Gilbert et al[8] and Robert[16].
- In addition, Nimblegen and Agilent microarray platform providers have evaluated and recommend Transplex WTA2[10-12].
- Procedures for integration of Transplex WTA2 and the popular microarray platform workflows, and for downstream qPCR application are provided on the Sigma-Aldrich website[20].
- Though not reported in the literature, Transplex WTA2 can serve as a vital "archiving" tool, storing RNA sequence information as stable cDNA for future alternative analytical methods.

Modification of Transplex-Amplified RNA for NGS Sequencing

The non-self complementary primer sequence is the linchpin of Transplex amplification, critical for efficient, positionally unbiased amplification of each transcript and comprehensive coverage of the transcriptome. The 5′ universal sequence spans 22 nucleotides, with the quasi-random 3′ region comprising 10 nucleotides (**Transplex WTA2 Workflow**). Because of the significant cost of labeled nucleotides and characteristic short read lengths for the Illumina and ABI platforms (**Table 1**), it is cost-effective to remove these primer sequences prior to NGS sequencing.

Sequencing Platform	Average Single-End Read	Average Paired-End Read
Illumina GA	72 bp	80 bp
Illumina GAIIx	150 bp	2 x 150 bp
Illumina HiSeq2000	35 bp	2 x 100 bp
ABI SOLiD 5500	75 bp	75 x 35 bp
ABI SOLiD 5500xi	75 bp	75 x 35 bp
Roche 454 GS Junior	400 bp	
Roche 454 GS FLX	400 bp	

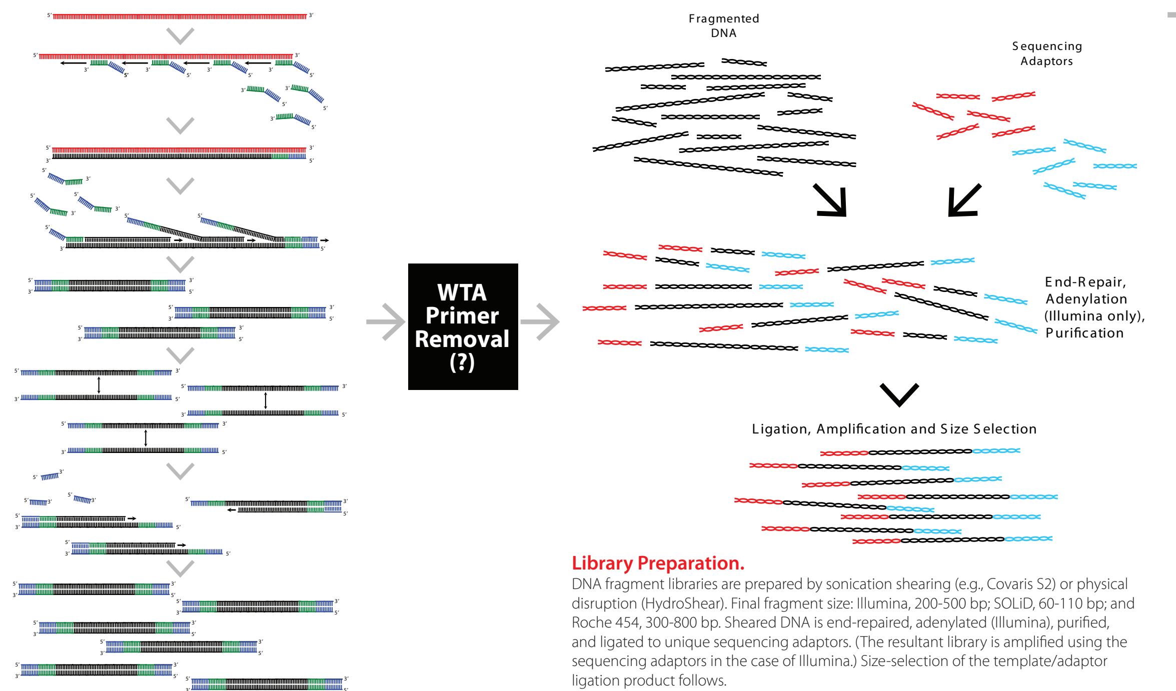
Table 1: NGS Sequencing Platform Read Lenghts[21,22]

Several enzymatic processes for primer sequence removal are under consideration.

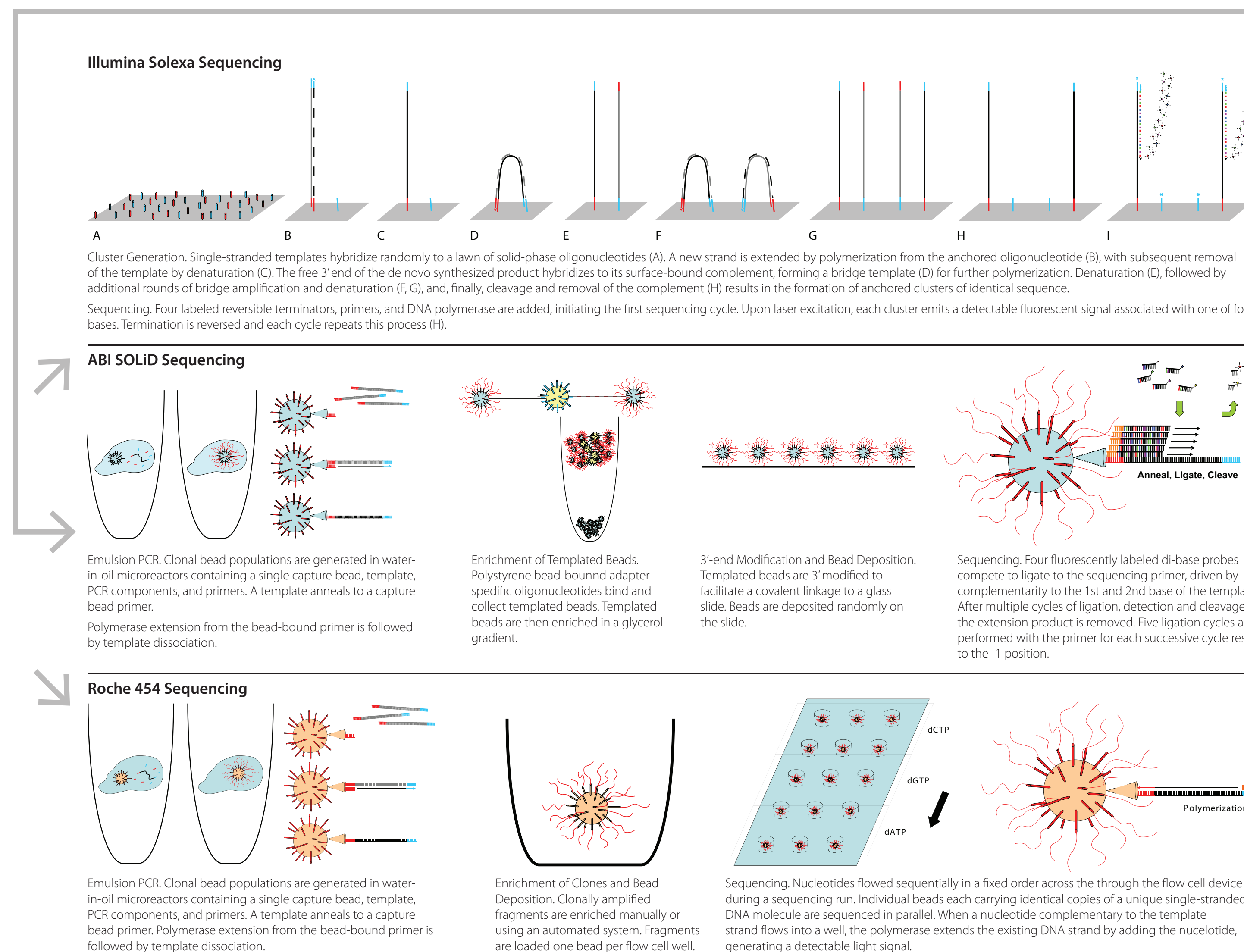
- Incorporation of a type II restriction site in the universal primer sequence. This would allow for cleavage of both universal and quasi-random primer sequence from the amplified cDNA prior to sequencing adaptor ligation (**Library Preparation**).
- Amplification incorporating modified dNTPs. Random incorporation of modified dNTPs during amplification is followed by nuclease digestion and 3′ end polishing (and dA tailing as required). This approach will statistically remove a majority of WTA library and amplification primer-derived sequences[23] prior to sequencing adaptor ligation (**Library Preparation**).

Transplex WTA2 Workflow

A cDNA strand is reverse transcribed from total RNA with the unique non-self complementary Transplexlibrary synthesis primer, comprised of a 3′ quasi-random region (**green**) and a defined universal 5′ end (**blue**) (A-C). Following RNase H degradation of the RNA template, strand-displacement polymerization generates a second cDNA strand (D, E). Additional transcription from this second stand also occurs at this step (not shown). Amplification of the double-strand cDNA library is then performed using a single universal primer (F-H).



Library Preparation. DNA fragment libraries are prepared by sonication shearing (e.g., Covaris S2) or physical disruption (HydroShear). Final fragment size: Illumina, 200-500 bp; SOLiD, 60-110 bp; and Roche 454, 300-800 bp. Sheared DNA is end-repaired, adenylated (Illumina), purified, and ligated to unique sequencing adaptors. (The resultant library is amplified using the sequencing adaptors in the case of Illumina.) Size-selection of the template/adaptor ligation product follows.



Illumina Solexa Sequencing. Cluster Generation. Single-stranded templates hybridize randomly to a lawn of solid-phase oligonucleotides (A). A new strand is extended by polymerization from the anchored oligonucleotide (B), with subsequent removal of the template by denaturation (C). The free 3′ end of the de novo synthesized product hybridizes to its surface-bound complement, forming a bridge template (D) for further polymerization. Denaturation (E), followed by additional rounds of bridge amplification and denaturation (F, G), and, finally, cleavage and removal of the complement (H) results in the formation of anchored clusters of identical sequence.

Sequencing. Four labeled reversible terminators, primers, and DNA polymerase are added, initiating the first sequencing cycle. Upon laser excitation, each cluster emits a detectable fluorescent signal associated with one of four bases. Termination is reversed and each cycle repeats this process (H).

ABI SOLiD Sequencing. Emulsion PCR. Clonal bead populations are generated in water-in-oil microreactors containing a single capture bead, template, PCR components, and primers. A template anneals to a capture bead primer.

Polymerase extension from the bead-bound primer is followed by template dissociation.

Enrichment of Templated Beads. Polystyrene bead-bounnd adapter-spedific oligonucleotides bind and collect templated beads. Templated beads are then enriched in a glycerol gradient.

3′-end Modification and Bead Deposition. Templated beads are 3′ modified to facilitate a covalent linkage to a glass slide. Beads are deposited randomly on the slide.

Sequencing. Four fluorescently labeled di-base probes compete to ligate to the sequencing primer, driven by complementarity to the 1st and 2nd base of the template. After multiple cycles of ligation, detection and cleavage, the extension product is removed. Five ligation cycles are performed with the primer for each successive cycle reset to the -1 position.

Roche 454 Sequencing. Emulsion PCR. Clonal bead populations are generated in water-in-oil microreactors containing a single capture bead, template, PCR components, and primers. A template anneals to a capture bead primer. Polymerase extension from the bead-bound primer is followed by template dissociation.

Enrichment of Clones and Bead Deposition. Clonally amplified fragments are enriched manually or using an automated system. Fragments are loaded one bead per flow cell well.

Sequencing. Nucleotides flowed sequentially in a fixed order across the through the flow cell device during a sequencing run. Individual beads each carrying identical copies of a unique single-stranded DNA molecule are sequenced in parallel. When a nucleotide complementary to the template strand flows into a well, the polymerase extends the existing DNA strand by adding the nucleotide, generating a detectable light signal.

An alternative to enzymatic manipulation of the Transplex WTA2-amplified library is the following.

- Sequencing oligonucleotide adaptors are ligated to the Transplex WTA2 amplicon (without removal of any primer sequence), as required by the respective sequencing platform workflow (**Library Preparation**).
- The platform sequencing primer would be replaced by an alternative primer complementary to the Transplex WTA2 universal amplification primer (**Sequencing: Illumina, Roche 454**).
- This approach will eliminate the amplification primer sequence from further analysis, but leave the quasi-random library synthesis sequence.

Discussion

The expectations of researchers are driving further development of methodologies and technology for single-cell analyses. Examples of RNA-Seq[24-26] and RIP-Seq[27] analyses (**RNA Immunoprecipitation**) are becoming more prevalent, in the study of disease and development, particularly in the assessment of oocyte and embryo development[28].

Sequencing Platform	Input Fragment Quantitiy
Illumina GAIIx	100 ng
ABI SOLiD 5500xi	500 ng
Roche 454 FLX	1 ug

Table 2: NGS Sequencing Platform Fragment Input Quantities[30-32]

Challenges for NGS Sequencing

- Cell expression profiling requires robust methods of RNA amplification for generation of sufficient starting material for deep sequencing[29] (**Table 2**).

Transplex WTA2 produces 3 to 5 ug of amplified product from 20 picograms of input total RNA (a single-cell equivalent quantity).

- Normalization between samples must account for expression level, gene length of the genes, and the complexity of the RNA population[33].

Conclusions

- Transplex Complete WTA2 kit has been demonstrated to be the method of choice for amplification of damaged and/or low quantities of RNA for microarray application.
- Efforts are underway to modify the Transplex WTA2 amplicon to allow for easy and efficient entry into workflows of the current popular deep sequencing platforms.

References

1. Heuermann, Ken and Brian Ward, Sigma-Aldrich Corporation, St. Louis, MO. *Whole Transcriptome Amplification. Posters and publications: Detection of Prostate Cancer Biomarkers by Expression Microarray Analysis of Transplex™ WTA2-Amplified FFPE Tissue RNA.* 2010 ABRF Annual Meeting. Sacramento, CA. Poster Abstract.http://www.sigmaaldrich.com/sigma-aldrich/areas-of-interest/life-science/molecular-biology/whole-genome-amplification/whole-transcriptome.html#posters
2. Heuermann, Ken and Brian Ward, Sigma-Aldrich Corporation, St. Louis, MO. *Whole Transcriptome Amplification. Posters and publications: Transplex™ Whole Transcriptome Amplification of RNA from Low Cell-Number Samples.* 2010 ABRF Annual Meeting Sacramento, CA. Poster Abstract. http://www.sigmaaldrich.com/sigma-aldrich/areas-of-interest/life-science/molecular-biology/whole-genome-amplification/whole-transcriptome.html#posters
3. Cadwel, Ken et al. *Virus-Plus-Susceptibility Gene Interaction Determines Crohn's Disease Gene Atg16L1 Phenotypes in Intestine.* 2010. Cell 141(7): 1135
4. Gonzalez-Roca, Eva et al. *Accurate Expression Profiling of Very Small Cell Populations.* 2010. PLoS ONE 5(12): e14418
5. Paik, Soonmyung, Samsung Cancer Research Institute, Samsung Medical Center, Seoul, Korea. *Interrogation of Gene Expression Using RNA Extracted from Formalin-Fixed Paraffin Embedded Tissue Blocks.* 2010. 4th World Congress, International Head and Neck Oncologic Society. Keynote Speaker
6. Flynn, James, Postdoctoral Fellow, Buck Institute for Age Research. Select Biosciences Single Cell Analysis Summit, La Jolla, CA. *Examination of Single Mouse Cardiomyocytes: Gene Expression Analysis and Mitochondrial Assay Development.* October 28-29, 2010. Select Biosciences Single Cell Analysis Summit, La Jolla, CA. Poster Abstract
7. Deal, Roger B. and Steven Henikoff. *The INTACT method for cell type–specific gene expression and chromatin profiling in Arabidopsis thaliana.* 2011. Nature Protocols 6, 56
8. Gilbert, Isabelle et al. *Providing a stable methodological basis for comparing transcript abundance of developing embryos using microarrays.* 2010. Molecular Human Reproduction 16(8): 601
9. Risk, Michael C. et al. *Differential Gene Expression in Benign Prostate Epithelium of Men with and without Prostate Cancer: Evidence for a Prostate Cancer Field Effect.* 2010. Clin Cancer Res 16: 5414
10. Skalitzky, Courtney and Mary-Anne Watt. *Technical Note: Evaluation of RNA amplification methods for NimbleGen Gene Expression microarray analysis.* 2010. http://www.nimblegen.com/products/lit/NG_GeneExp_TechNote_RNAAmplification_101310.pdf
11. Bergstrom-Lucas, A. and Gary Lin. 2009. *Application Note: Gene Expression Microarray Analysis of Archival FFPE Samples.* http://www.chem.agilent.com/Library/applications/5990-3917en_hi.pdf
12. Bergstrom-Lucas, A. et al, Agilent Technologies, Santa Clara, CA. *A new workflow for sensitive and accurate microarray gene expression profiling from formalin-fixed, paraffin-embedded (FFPE) tissues.* 2009. Proc Am Assoc Cancer Res; 2009 Apr 18-22: Abstract #5208. The 100th AACR Annual Meeting; Denver, CO. http://www.aacrmeetingabstracts.org/cgi/content/meeting_abstract/2009/2_Annual_Meeting/5208?maxtoshow=&hits=10&RESULTFORMAT=1&author1=Anne+Bergstrom+Lucas&author2=Gary+lin&andorexacttitle=and&andorexacttitleabs=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&fdate=1/1/2004&tdate=12/31/2010&resourcetype=HWCIT
13. Laxman, Bharathi et al. *A First-Generation Multiplex Biomarker Analysis of Urine for the Early Detection of Prostate Cancer.* 2008. Cancer Res 68: 645
14. Ward; Brian and Heuermann; Kenneth E. *Degenerate Oligonucleotides and Their Uses.* April 16, 2009. US Patent Application 20090099040
15. *Personal Communication:* Richard Head, Associate Research Fellow, Inflammation and Immunology Research Unit, Pfizer Global Research and Development, St. Louis, MO 63017, USA
16. Robert, Claude. *Microarray analysis of gene expression during early development: a cautionary overview.* 2010. Reproduction 140: 787
17. Van Gelder, R.N. et al. *Amplified RNA synthesized from limited quantities of heterogeneous cDNA.* 1990. Proc. Natl. Acad. Sci. USA 87: 1663-1667
18. Kamberov, E. et al. *Amplification and analysis of whole genome and whole transcriptome libraries generated by a DNA polymerization process*. May 18, 2010. US Patent 7,718,403
19. Ward, Brian, Associate Fellow, Sigma Aldrich Corporation. *The Quest for Faithful Amplification of Degraded and Microscopic Quantity RNAs.* October 28-29, 2010. Select Biosciences Microarray World Congress, La Jolla, CA. Technology Spotlight
20. *Whole Transcriptome Amplification. NEW! Labeling Protocols for Microarrays.* http://www.sigmaaldrich.com/sigma-aldrich/areas-of-interest/life-science/molecular-biology/whole-genome-amplification/whole-transcriptome.html#posters
21. Karow, Julia. *Survey: Illumina, SOLiD, and 454 Gains Ground in Research Labs; Most Users Mull Additional Purchases.* January 19, 2010. Gemomeweb In Sequence http://www.genomeweb.com/sequencing/survey-illumina-solid-and-454-gain-ground-research-labs-most-users-mull-addition
22. *Next Generation Sequencing Spreadsheet.* December 2010. http://ngsbuzz.blogspot.com/ (To view spreadsheet, Safari or Mozilla FireFox web browser may be required.)
23. Wang, Z. et al. *Overproduction and characterization of the uracil-DNA glycosylase inhibitor of bacteriophage PBS2.* 1991. Gene 99: 31.
24. Lozano, Gloria et al. *Populations of Genomic RNAs Devoted to the Replication or Spread of a Bipartite Plant Virus Differ in Genetic Structure.* 2009. J. Virol. 83: 12973
25. Strable, Josh and Michael J. Scanlon. *Maize (Zea mays): A Model Organism for Basic and Applied Research in Plant.* 2009. Cold Spring Harb Protoc 2009:132
26. Laurie, Nikia et al. *Changes in Retinoblastoma Cell Adhesion Associated with Optic Nerve Invasion.* 2009. Mol. Cell. Biol. 29: 6268
27. Matkovich, et al. *RISC RNA Sequencing for Context-Specific Identification of In Vivo MicroRNA Targets.* 2011. Circ. Res. 108:18
28. Seli, Emre et al. *OMICS in Assisted Reproduction: Possibilities and Pitfalls.* 2010. Molecular Human Reproduction 16(8): 513
29. Wang, Z. et al. *RNA-Seq: a revolutionary tool for Transcriptomics.* 2009. Nature Reviews Genetics 10: 57
30. *GAIIx Specification Sheet: Illumina® Sequencing* http://www.illumina.com/documents/products/datasheets/datasheet_genome_analyzerIIx.pdf
31. *Agilent SureSelect Target Enrichment System for the Applied Biosystems SOLiD System.* http://www.genomics.agilent.com/files/Manual/G3360-90021_SureSelect_SOLiDIndexing_1.1.pdf
32. *Transcriptome sequencing with the Genome Sequencer FLX system* http://www.454.com/downloads/protocols/Transcriptome_NatureMethods.pdf
33. Robinson, MD and A. Oshlack. *A scaling normalization method for differential expression analysis of RNA-Seq data.* 2010. Genome Biology 11: R25

To contact author, please email ken.heuermann@sial.com