Genomic Analysis of Dissociated Tumor Cells

Shawn P. Fahl, PhD

ABSTRACT

Next-generation deep sequencing is a powerful tool to understand the global genomic mutations and transcriptional alterations associated with complex diseases, such as cancer and autoimmunity. Dissociated tumor cells, or DTCs, are cryopreserved single cell suspensions that consist of the complete cellular composition of the tumor microenvironment. DTCs are amenable to numerous downstream applications. To investigate the utility of DTCs for genomic applications, we performed whole exome sequencing (WES) and whole transcriptome sequencing (RNASeq) on eight DTC samples across four indications. High quality DNA and RNA was isolated from each DTC sample at quantities suitable for downstream WES and RNASeq analysis. Whole exome analysis revealed numerous genomic variants, including previously described pathogenic alleles in clinically-relevant oncogenes. Whole transcriptome analysis revealed differential gene expression across each of the samples, although the samples did cluster based on their tissue of origin. The gene expression data was then further subdivided on specific gene subsets to highlight cellular functionalities such as T cell responsiveness. This study confirms the suitability of cryopreserved tumor derived single-cell suspensions for downstream genomic analysis to identify known and novel genomic alterations and gene expression changes.

SAMPLE SET AND WORKFLOW

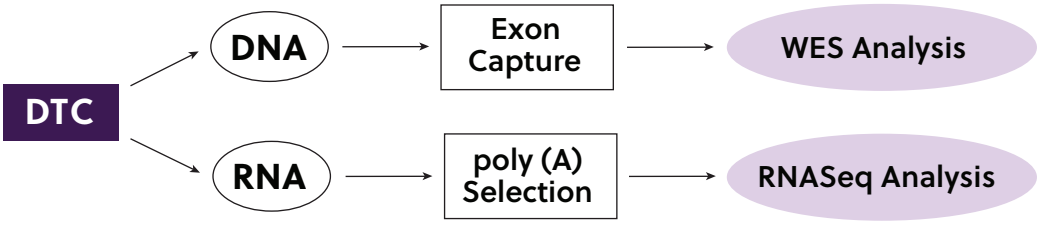
Dissociated tumor cells (DTCs) are cryopreserved single-cell suspensions generated from solid tumors via mechanical and enzymatic digestions. These samples contain all the cellular constituents of the tumor microenvironment and are ready for numerous downstream applications, including flow cytometry, cell culture, and next-generation sequencing. To further validate the utility of DTCs in next-generation sequencing, eight unique DTC samples across four different indications - melanoma, lung cancer, ovarian cancer, and bladder cancer - were selected for whole exome and whole transcriptome analysis (Table 1). Samples were screened by flow cytometry to ensure a high percentage of tumor cells within the samples. In general, melanoma, ovarian cancer, and bladder cancer DTCs have a higher percentage of tumor cells than lung cancer DTCs[1]; therefore, the lung cancer samples selected for analysis show slightly lower tumor cell percentages comparatively. DNA and RNA was isolated from cryopreserved DTCs from each patient for whole exome and whole transcriptome analysis, respectively (Figure 1). On average, a yield of 4 µg DNA and 2.5 µg RNA, with an average RIN value of 8.8, was observed per 1 x 10^6 cells

TABLE 1. Sample Set for Genomic Analysis.

Patient	Indication	Stage	Tumor	Immune
110003345	Lung	III-A	37.90%	50.80%
110003620	Lung	II-A	56.30%	40.50%
110003616	Melanoma	III-C	75.10%	23.30%
110005738	Melanoma	III-C	86.00%	4.25%
110003075	Ovarian	I	97.80%	1.74%
110002936	Ovarian	III-C	69.90%	26.00%
110006040	Bladder	II	72.70%	19.70%
110003631	Bladder	I	71.70%	22.90%

Eight DTC samples across 4 indications were analyzed by flow cytometry to determine relative contribution of tumor and immune cells and selected for genomic analysis.

FIGURE 1. Workflow for Genomic Analysis of DTCs.



DNA and RNA was isolated from each patient DTC sample. DNA was analyzed by whole exome sequencing, while RNA was analyzed by RNASeq.

WHOLE EXOME ANALYSIS

The exome represents all the protein-coding regions (or exons) of the genome, and most known disease-causing genetic mutations occur within these exons. Whole exome sequencing provides an unbiased method to evaluate genetic mutations across all exons, including known clinically-relevant oncogenes as well as novel, previously unknown oncogenic proteins. Across the 8 patient samples analyzed, ~300,000 genetic variants were identified per sample, including both single nucleotide polymorphisms (SNPs), insertions, and deletions (INDELs) (Table 2). Lung DTC sample from patient 110003345 and melanoma DTC sample from patient 110003616 were further analyzed for the clinically relevant oncogenes for each indication. For patient 110003345, numerous variants were observed in the EGFR, KRAS, and NRAS loci, including several that have been reported to correlate with disease[2-5] (Table 3). In melanoma patient 110003616, a pathogenic variant at amino acid 600 of the BRAF gene was also identified[6,7] (Table 4).

TABLE 2. Total Variants from Whole Exome Sequencing Analysis.

Patient	Indication	Variants	SNPs	INDELs
110003345	Lung	306101	257557	48544
110003620	Lung	278780	234989	43791
110003616	Melanoma	278227	235337	42890
110005738	Melanoma	264421	222973	41448
110003075	Ovarian	306066	256017	50049
110002936	Ovarian	340699	288397	52302
110006040	Bladder	307065	262345	44720
110003631	Bladder	317221	268901	48320

DNA was isolated from 8 unique DTC samples. Whole exome libraries were prepared using the NimbleGen SeqCap EZ Exome Library v3.0 and sequenced on a NovaSeq. Data was analyzed using DRAGEN.

TABLE 3. Analysis of Genetic Variants in DTC Lung Sample 110003345.

Patient: 110003345			
Indication: Lung			
ID	**Reference**	**Alternate**	**Gene**
RS1558544	A	T	EGFR
RS1140475	T	C	EGFR
RS1050171	G	A	EGFR
RS2692456	G	A	EGFR
RS4947986	G	A	EGFR
RS759162	C	T	EGFR
RS6970262	A	G	EGFR
RS34723095	C	CAG	EGFR
RS10241326	T	C	EGFR
RS2293348	C	T	EGFR
RS4947987	G	C	EGFR
RS712831	T	C	EGFR
RS3752651	T	C	EGFR
RS13222549	G	C	EGFR
RS6964705	C	A	EGFR
RS148883365	A	AAC	EGFR
RS55678943	G	A	EGFR
RS2472520	G	C	EGFR
RS371572654	A	G	EGFR
RS141162488	TACAC	T	EGFR
.	C	A	EGFR
RS1076452	G	A	EGFR
RS55920948	G	GACAC	EGFR
RS4597149	T	C	KRAS
RS4285970	G	A	KRAS
RS57698689	CTT	C	KRAS
RS7973623	G	A	KRAS
RS71065923	T	TA	KRAS
RS7973450	A	G	KRAS
RS7960917	T	C	KRAS
.	G	C	KRAS
RS1137282	A	G	KRAS
RS11836509	T	G	KRAS
RS34176876	TA	T	KRAS
RS712	A	C	KRAS
RS1137188	G	A	KRAS
RS8720	T	C	KRAS
RS1137196	T	G	KRAS
RS12245	A	T	KRAS
RS1137189	A	T	KRAS
RS9266	A	G	KRAS
.	T	TAA	KRAS
RS14804	G	A	NRAS
.	T	C	NRAS
RS528488691	T	C	NRAS

SNPs and INDELs were analyzed in the clinically relevant oncogenes EGFR, KRAS, and NRAS. Variants that have been previously implicated in disease are highlighted in red.

TABLE 4. Analysis of Genetic Variants in DTC Melanoma Sample 110003616.

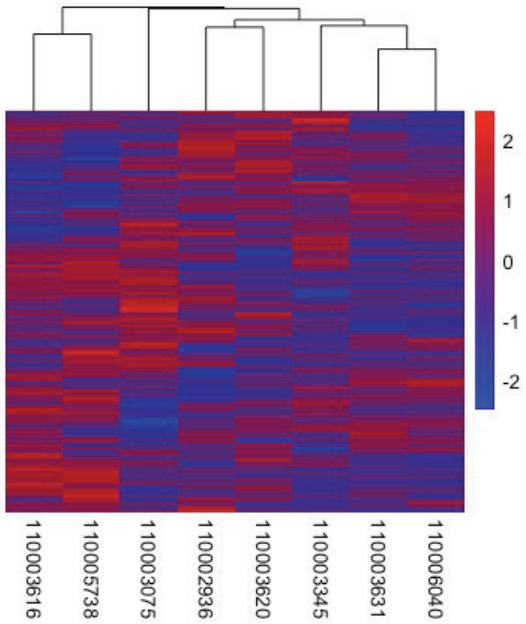
Patient: 110003616			
Indication: Melanoma			
ID	**Reference**	**Alternate**	**Gene**
RS3829814	A	G	BRAF
RS60814637	G	GA	BRAF
RS113488022	A	T	BRAF
RS750726638	CAA	C	BRAF
RS1733826	T	C	BRAF
.	CA	C,CAA	BRAF
RS969273	G	A	NRAS
RS7549358	G	C	NRAS

SNPs and INDELs were analyzed in the clinically relevant oncogenes BRAF and NRAS. Variants that have been previously implicated in disease are highlighted in red.

WHOLE TRANSCRIPTOME ANALYSIS

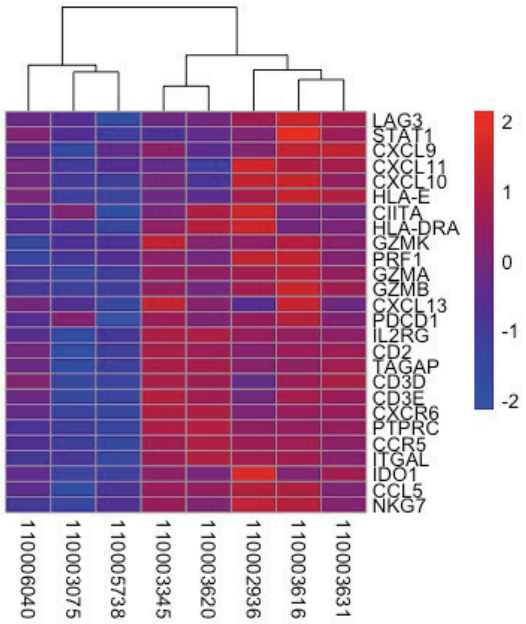
The transcriptome represents all the RNA molecules that are expressed, and the level to which they are expressed, within the cell. Unlike the genome, which is relatively static within an organism, the transcriptome varies widely across cell types and is highly influenced by the external environment. Analysis of the all of the genes expressed in the 8 patient samples revealed significant differential expression of genes across all samples (Figure 2). However, the samples did segregate by their tissue of origin when the entire transcriptome of each sample was incorporated, with melanoma clustering away from the three epithelial-based indications. The ovarian DTC sample from patient 110003075 did segregate away from ovarian DTC sample 110002936, consistent with the small percentage of immune cell infiltrate observed in sample 110003075. The transcriptome data set can be further subdivided to analyze distinct subsets of genes, such as T cell signature genes that have been utilized in other studies[8] (Figure 3). When this subset of genes was analyzed, three patient samples (110006040, 110003075, and 110005738) had substantially lower expression of T cell signature genes than the other samples analyzed, consistent with the low percentage of immune cells present in samples 110003075 and 110005738. Patient sample 110006040, on the other hand, did have a significant population of immune cells, suggesting that these cells have be hypofunctional based on transcriptional analysis.

FIGURE 2. Whole Transcriptome Analysis of DTCs.



RNA was isolated from 8 unique DTC samples. Standard polyA RNASeq libraries were prepared and sequenced on a NovaSeq. Reads were mapped using TopHat and analyzed using Cufflinks. Data are represented as log2 FPKM values.

FIGURE 3. T Cell Signature from Whole Transcriptome Analysis.



RNASeq data was analyzed for T Cell Signature genes[8]. Data are represented as log2 FPKM values.

REFERENCES

1. "Large-Scale Flow Cytometry Analysis of Tumor Tissues" https://www.dls.com/large-scale-flow-whitepaper
2. Fung, C., *et al.* "Identification of epidermal growth factor receptor (EGFR) genetic variants that modify risk for head and neck squamous cell carcinoma." *Cancer Lett.* 2005 Feb 28;357(2):549-56
3. Hasheminasab, S.M., *et al.* "High-throughput screening identified inherited genetic variations in the EGFR pathway contributing to skin toxicity of EGFR inhibitors." *Pharmacogenomics.* 2015;16(14):1605-19
4. Li, B., *et al.* "Effect of epidermal growth factor receptor gene polymorphisms on prognosis in glioma patients." *Oncotarget.* 2016 Sep 27;7(39):63054-64
5. Egeli, U., *et al.* "Impact of 3'UTR variation patterns of the KRAS gene on the aggressiveness of pancreatobiliary tumors in the KRAS G13D mutation in a Turkish population." *Pancreatology.* 2016 Jul-Aug;16(4):677-86.
6. Kumar, R., *et al.* "BRAF mutations in metastatic melanoma: a possible association with clinical outcome." *Clin Cancer Research.* 2003 Aug 15;9(9):3362-8
7. MacConaill, L.E., *et al.* "Prospective enterprise-level molecular genotyping of a cohort of cancer patients." *J Mol Diagn.* 2014 Nov;16(6):660-72
8. Ayers, M., *et al.* "IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade." *J Clin Invest.* 2017 Aug 1;127(8):2930-40