Identification and temporal monitoring of breast cancer-associated T cell receptors with high throughput sequencing

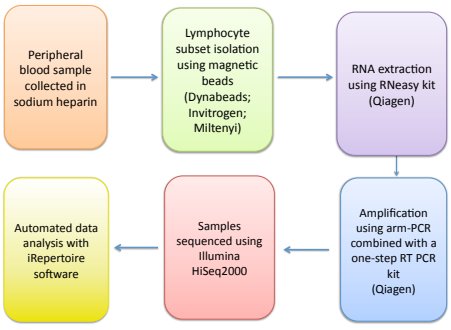
Jian Han[1], Chunlin Wang[1], Qunying Yang[1], Catherine M. Sanders[1], Miranda Byrne-Steele[1], Enders Ng[2], Chris L.P. Wong[3], Ava Kwong[2], Chiung-Mei Lu[1]

[1]iRepertoire Inc., HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA, [2]Breast Surgery, Hong Kong University, Hong Kong, China
[3]Deparment of Pathology, Hong Kong Sanatorium and Hospital, Hong Kong, China

Abstract

Tumor-specific antigens may trigger an immune response that leads to T lymphocytes infiltrating the tumor tissue. We have developed a method to study the immune repertoire of a sample by utilizing a patented multiplex PCR amplification strategy, arm-PCR (Patent No. 7,999,092), coupled with high throughput sequencing (Wang et al., PNAS 2010). Using this method, we sequenced CDR3 fragments amplified from cancer tissue and normal tissue surrounding the cancer sites, which were surgically removed from a patient during operation. The patient's sorted peripheral blood was also examined at three months, six months, and one year after surgery. Dominant T cell clones with specific CDR3 sequences were identified from the breast cancer tissue. Some of these same clones were found expressed at high levels in the nearby normal tissue and peripheral CD8⁺ cells. After treatment, dynamic changes in these cancer-associated clones were apparent, demonstrating the capability of the current technology to identify specific T cells associated with a patient's cancer tissue. These specific T cells can serve as personalized biomarkers for prognosis, treatment evaluations, and early detection of recurrence. They can also be used to develop personalized treatment strategies. Currently, this study has been extended to examine additional breast cancer patients.

Methods



Results

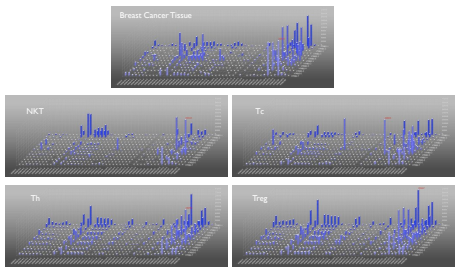


Figure 1. The distribution of domain usage for TCR beta chain in breast cancer-associated T cells, NK T cells, Tc, Th, and Treg cells sorted from peripheral blood. The height of bar shows the frequency of V-J combination usage.

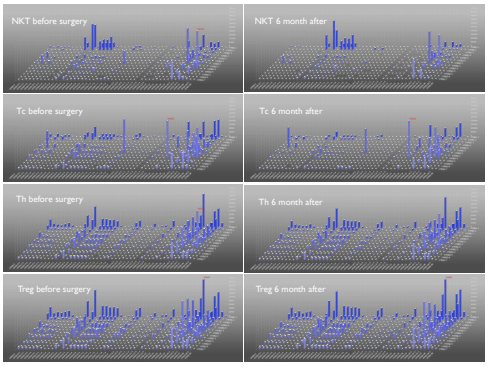


Figure 2. The overall patient repertoire is relatively stable, as demonstrated in the repertoires before and 6 months after the surgery.

References

Wang, Chunlin *et al.* High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *PNAS*. 107,1518-1523.(2010).

Results

CDR3	NKT before	NKT 6 mon	CDR3	CyT before	CyT 6 mon
SARTAPSYNEQF	334154	427446	ASSSRLAGGTDTQY	205516	373620
ASSFPSGRADTQY	295191	50098	ASSQGGTGPVPNQPQH	194818	317287
ASSQGGTGPVPNQPQH	117437	19395	ASSFPSGRADTQY	111683	14274
ASSSRLAGGTETQY	100388	3224	ASSSRLAGGTETQY	105202	179859
ASSFPLTGAKNIQY	77709	119305	ASSVPGEFSNEQF	61851	47344
ASSSRLAGGTDTQY	63846	9156	ASSPQGGRLNTEAF	41491	44482
ASSFTRGARYT	29327	21073	SARTAPSYNEQF	35072	28101
ASSPQGGRLNTEAF	27510		ASSANRGTEAF	34919	24063
ASSVPGEFSNEQF	27190		SARDPAGNFDEQY	34066	46527
ASSLAGYFYEQY	24428		ASSHRTGLAYEQY	33131	14751
SARVGDRVSSSYNSPLH	24139	34196	ASSLSHRNEQY	27002	20621
SAKGLWDASYGYT	23434		SAKGLWDASYGYT	25027	14904
SARLADNEQF	19205		ATSVTERREQY	19970	29793
ASKGDPYETGELF	19166		ASSLAGYFYEQY	19705	
SASSGTGVDNEQF	17752		SARLADNEQF	19558	21492
ASSLRGHDSPLH	15572	15431	SASSGTGVDNEQF	19179	17332
SAQKVLELNTEAF	13721		ASKGDPYETGELF	17741	9378
SAGTAEPYEQY	12908	3983	SVKREGNEQF	16961	18933
SVKREGNEQF	11682		ASSFPLTGAKNIQY	16170	8258
ASSHRTGLAYEQY	10397		ASSEWGVQYEQY	13123	4796

CDR3	HpT before	HpT 6 mon	CDR3	RgT before	RgT 6 mon
ASSFPSGRADTQY	162602	49309	SAPSLSGSTDTQY	5346	4871
SARTAPSYNEQF	24869	32296	ASSGQMNTEAF	3767	3323
SAVGHNNQPQH	12521	9116	ASIEGGYSNQPQH	1814	3280
ASSGQMNTEAF	11510	16394	ASSFPSGRADTQY	8187	3069
ASSFPLTGAKNIQY	9703	22677	SARDFGRQGPMLGQF	2059	2808
SARVGDRVSSSYNSPLH	8601	19429	ASSFMGGQETQY	1942	2758
ASSSFDMNTEAF	3262	4506	SAVGHNNQPQH	11572	2382
ASSLDNEQF	3215	3900	ASTPVGSSYNEQF		2314
ASSESRGTNEKLF	3185	5467	SANSLVSADTQY	4092	1863
SAQNQETQY	2752		ASSEDLNQPQH	2214	1572
ASSYHGDEQF	1890		SARVGDRVSSSYNSPLH		1343
ASSEDLNQPQH	1817	2056	ASSRDSNTEAF		1242
SANSLVSADTQY	1753	1506	ASRNSNQPQH	1031	1210
ASRNSNQPQH	1691	3025	SAPAEATGSGNTIY	1252	1104
ASSSRLAGGTDTQY	1675	1263	ASSGQLNTEAF	1648	1075
ASSPKESNEQF	1534	1781	SAQNQETQY	3515	
ASSGQLNTEAF	1528	1115	ASSYHGDEQF	1967	
ASGGLTEGLNIQY	1481	2418	ASSYQSNEQF	1078	
ASSFEVNTGELF	1410		SARTAPSYNEQF	1055	
ASSYLRENTEAF	1209		ASRAGTFTEAF	1025	

Table 1. Changes in expression level of the top 20 cancer related CDR3 in each subsets of T cells. After normalizing, those with more than two fold changes were highlighted. Significant number of NK T cells were down regulated.

Conclusions

- Combining arm-PCR with high-throughput sequencing, we are able to identify specific T cells associated with a patient's cancer tissue in both the tumor and the peripheral blood.

- Specific T cell CDR3 sequences identified in cancer patients may serve as personalized biomarkers for prognosis, treatment evaluation, and early detection of recurrent tumors.

- Analysis of additional breast cancer tissue samples together with peripheral blood repertoires has allowed identification of shared CDR3 sequences among patients, indicating the potential existence of a group of disease-specific T cells sharing CDR3 sequences.

- Further studies are needed to identify and verify specific CDR3 sequences that may be unique to breast cancer and serve as biomarkers for disease.