Lymphocytes play a critical role in antitumor immune responses. They are directly targeted by some therapies, and the composition and spatial organization of intratumor T-cell populations is prognostic in some cancer types. A better understanding of lymphocyte population dynamics over the course of disease and in response to therapy is urgently needed to guide therapy decisions and to develop new therapy targets. Deep sequencing of the repertoire of antigen receptor–encoding genes expressed in a lymphocyte population has become a widely used approach for profiling the population’s immune status. Lymphocyte antigen receptor repertoire deep sequencing data can be used to assess the clonal richness and diversity of lymphocyte populations; to track clone members over time, between tissues, and across lymphocyte subsets; to detect clonal expansion; and to detect the recruitment of new clones into a tissue. Repertoire sequencing is thus a critical complement to other methods of lymphocyte and immune profiling in cancer. This review describes the current state of knowledge based on repertoire sequencing studies conducted on human cancer patients, with a focus on studies of the T-cell receptor beta chain locus. The review then outlines important questions left unanswered and suggests future directions for the field.

In the decades since, we have acquired a better understanding of the complex interactions and mutual influence between tumor cells and the host immune system (10, 11). While the immune system can recognize and destroy tumor cells, tumors can evade and suppress immune responses. In addition, selection pressure from antitumor immune responses contributes to the evolution of immune resistant tumors. This improved understanding led to the immunomodulating hypothesis, an extension of the cancer immunosurveillance hypothesis that includes three phases: elimination of neoplastic cells by the immune system; equilibrium, in which the immune system controls but does not eliminate cancer; and escape, in which cancer progresses to clinically detectable disease (5, 12). Immunotherapy was beautifully demonstrated in ref. 13, which showed that, in early stage, untreated non–small-cell lung cancer, selection pressure from antitumor immune responses resulted in multiple routes of tumor immune evasion. Tumors with high levels of immune infiltrates exhibited neoantigen depletion (e.g., via copy number loss and reduced transcription and disrupted antigen presentation (e.g., via loss of human leukocyte antigen heterozygosity), while tumors with low levels of immune infiltrates did not (13). It was further found that the extent of immune evasion was prognostic (13).

Although all arms of the immune system participate in antitumor immune responses, multiple lines of evidence support a critical role for lymphocytes, particularly cytotoxic T lymphocytes (6, 9, 14, 15). In humans, there are two primary lines of evidence: (i) an association between survival in patients with cancer and the location, density, type, and functional orientation of tumor-infiltrating lymphocytes (TIL) and (ii) the effectiveness of immunotherapies that leverage natural, antitumor T-cell responses (1, 5–7, 10). Regarding (i), multiple studies have found that high densities of CD3+ T cells, CD8+ cytotoxic T cells, or CD45RO+ memory T cells, especially at the tumor center or invasive margin, are associated with increased progression-free and overall survival across multiple cancer types (16–24). Regarding (ii), evidence comes from adoptive cell transfer (ACT) therapy with natural, autologous T cells and checkpoint blockade therapy. In ACT therapy with natural, autologous T cells, a patient is repopulated with his or her own antitumor T cells that have been selected, activated, and
expanded ex vivo. The transferred cells traffic to the tumor and metastatic sites where they kill tumor cells, often resulting in tumor regression and improved survival (25–30). In some cases, complete response with long-term follow up has been observed (27). Checkpoint blockade therapy involves the administration of mAbs that interfere with signals that inhibit T-cell activation and expansion, signals utilized by tumor cells and by suppressive immune cells recruited by tumors. Checkpoint blockade therapy, such as with antibodies against cytotoxic T lymphocyte–associated antigen 4 (CTLA-4), programmed death 1 (PD-1), or its ligands (PD-L1, PD-L2), has resulted in tumor regression and/or improved survival for some patients (31–36).

**Lymphocyte antigen receptors, a.k.a. adaptive immune receptors**

B and T lymphocytes are the primary actors of the adaptive immune system, the subsystem of the immune system characterized by (i) responses that are highly specific for a particular target (antigen) and (ii) immune memory, the phenomenon whereby, on exposure to a particular antigen, the system acquires the ability to respond rapidly and with increased effectiveness on subsequent exposures to the same antigen. The high specificity of B and T lymphocyte responses is a consequence of the fact that the genes encoding their cell surface antigen receptors (adaptive immune receptors, AIR) are somatically generated in developing lymphocytes through a process that creates essentially unique gene sequences in each cell. Antibodies are a soluble form of the B lymphocyte antigen receptor and are encoded by the same gene. The somatic generation of AIR genes results in a tremendous diversity of the B lymphocyte antigen receptor and are encoded by the same gene. The somatic generation of AIR genes results in a tremendous diversity of AIR repertoires (AIRR) is critical for providing individuals with the ability to respond to the tremendous variety of antigens encountered over one's lifetime.

Lymphocytes constantly sample the antigens in their environment and proliferate when they encounter their cognate antigen (the antigen for which their antigen receptor has specificity), thereby forming a clone of cells expressing identical receptors. This clonal expansion, together with lymphocyte migration, birth, and death, means that the clonal composition of an AIRR evolves over time and reflects the history and current state of the population’s antigen exposures. Lymphocyte populations are also dynamic with respect to the types of lymphocytes composing them. Upon contact with their cognate antigen, lymphocytes become activated and acquire the ability to exert their effector function. For example, CD8+ cytotoxic T cells, once activated, can kill cells expressing their antigen receptor’s cognate antigen. In the face of chronic antigen stimulation, activated lymphocytes may become exhausted and lose their ability to exert their effector function. Activated lymphocytes may also differentiate to become memory cells, which are long-lived and capable of rapid response on subsequent encounter with their cognate antigen.

Because the genes encoding AIR are essentially unique to a lymphocyte clone, they serve as a tag for members of the clone. Thus, AIRR sequencing data can be used to: track clone members over time, between tissues, and across lymphocyte subsets; detect clonal expansion (i.e., a local, antigen-specific immune response); and detect the recruitment of new clones. In addition, the data can be used to estimate repertoire-level statistics designed to capture the number of distinct clones in a repertoire and the extent to which they have undergone clonal expansion. The most common statistics (richness, diversity, and clonality) correspond to concepts borrowed from community ecology (43). Richness is simply the number of different AIR in a repertoire. Diversity takes into account both the number of AIR in a repertoire and their relative abundances. For example, the most commonly used diversity measure is the Shannon diversity index, typically calculated as

\[ H = -\sum_{i=1}^{N} p_i \ln p_i, \]

where \( i \) indexes the distinct AIR in a repertoire, \( N \) is the total number of distinct AIR in the repertoire, and \( p_i \) is the relative abundance of the \( i^{th} \) AIR in the repertoire. Two repertoires with the same number, \( N \), of distinct AIR could have very different diversity values if, in one, all the clones were of similar abundance, whereas, in the other, some of the clones had undergone clonal expansion and therefore had relative abundances much larger than that of the other clones. For a given \( N \), the maximally diverse repertoire is the one for which the AIR are uniformly distributed (\( p_i \) are all equal). Estimates of clonality capture the extent to which clonal expansion has shifted the distribution of relative abundances away from uniform. A common measure of clonality is \( 1 - H/H_{\max} \) (43). The soundness of richness, diversity, and clonality estimates depends on the number of lymphocytes sampled and on the soundness of the composite clonal relative abundance estimates. Thus, as elaborated below in IMPLICATIONS, the reliability of repertoire-level statistics depends heavily on the sequencing protocol used. Nevertheless, sequencing the AIR genes of a lymphocyte population can be a powerful tool for assessing and predicting adaptive immune status, and it has become a critical complement to other methods of lymphocyte, and immune, profiling in cancer.

**AIR Repertoires in Cancer – Current State of the Field**

The application of deep AIRR sequencing to characterize lymphocyte responses in cancer is now widespread (43, 44). The results so far indicate tremendous potential for AIRR profiling not only to advance our understanding of the complex interactions between cancer, the immune system, and therapy, but also to make significant contributions in the clinic via novel approaches to diagnosis, prognosis, therapeutic decision making, and the development of new therapeutic targets. Although many interesting and important AIRR profiling studies are conducted using animal models, due to space constraints, this review is focused on AIRR sequencing in human patients with cancer. In addition, this review focuses on studies of the T-cell receptor (TCR) beta locus, as the vast majority of studies to date have focused on T-cell responses. The application of deep AIRR sequencing to characterize lymphocyte responses in cancer is now widespread (43, 44). The results so far indicate tremendous potential for AIRR profiling not only to advance our understanding of the complex interactions between cancer, the immune system, and therapy, but also to make significant contributions in the clinic via novel approaches to diagnosis, prognosis, therapeutic decision making, and the development of new therapeutic targets. Although many interesting and important AIRR profiling studies are conducted using animal models, due to space constraints, this review is focused on AIRR sequencing in human patients with cancer. In addition, this review focuses on studies of the T-cell receptor (TCR) beta locus, as the vast majority of studies to date have focused on this locus. The basic findings are described in the current section, while the implications of the findings when jointly interpreted, along with unanswered questions, are discussed in the following section.

**Characterizing AIRR in patients with cancer**

**Tumor-infiltrating T-cell repertoires**

AIRR profiling studies across multiple cancer types have provided evidence for clonal expansion within tumors (45–53). The degree of clonal expansion typically exceeds that in patient-matched, adjacent healthy tissue (45, 49–51), although there are exceptions (54, 55). These results are indicative of local, antigen-specific T-cell responses. Studies examining specific T-cell subsets have found that the expansion may be concentrated within the exhausted subset of CD8+ TILs (e.g., CD8+PD-1+) or the regulatory T-cell subset of TILs (48, 50, 53, 56). Indeed, clonality within the exhausted subset may be substantial. One study in melanoma found that the 30 most abundant clones accounted for 60% to 80% of the CD8+PD-1+
population but only 28% to 37% of the CD8⁺ PD-1⁺ population (48). Another study in non–small-cell lung cancer found that approximately 33% of TCR RNA could be attributed to expanded clones in the PD-1⁺ population, while only approximately 16% corresponded to expanded clones in the PD-1⁺ population (53). In both studies, individual clones included members from both subsets (53, 56). Studies in melanoma have also demonstrated that the most abundant CD8⁺ PD-1⁺ clones have tumor reactivity (48, 56). In one study, 61% of CD8⁺ PD-1⁺ TILs recognized autologous tumor, while only 7% of CD8⁺ PD-1⁻ TILs did (48).

Further studies comparing tumor and adjacent healthy tissue found larger numbers of unique TCR sequences in tumors, suggesting increased recruitment of T cells into tumors (45, 49, 57), but others have found the opposite (51, 58). Most studies have found identical TCR sequences in both the TILs and adjacent healthy tissue repertoires, suggesting trafficking of T cells between the sites (45, 48, 51, 53, 54). The degree of overlap varied considerably, however, both between patients and between cancer types.

**Spatial heterogeneity of the TILs repertoire**

The distribution of lymphocyte clones and subsets throughout lesions, as well as between tumors and metastatic sites, may have considerable heterogeneity. In renal cell carcinoma, it was shown via ultra-deep aIRR sequencing on multiple tumor regions, that few TCR sequences are found in all tumor regions and most T cell clones are spatially separated (46). Similarly, in non–small-cell lung cancer, highly abundant TCR sequences, presumably derived from expanded clones, could be classified as ubiquitous or regional, depending upon their spatial distribution (53). In both studies, for clones/sequences present in multiple tumor regions, the frequencies were poorly correlated across regions (46, 53). In lung adenocarcinoma, substantial intratumor heterogeneity was observed in T-cell density, clonality, and clonal distribution, with most clones restricted to a single tumor region (59). In a single patient with bilateral breast cancer, there was only 10% overlap in the TCR sequences between the tumors (60).

In contrast, an ovarian cancer study found that T-cell clones were generally spatially homogeneous in distribution throughout metastatic lesions and only slightly less so throughout tumors (61). There was, however, considerable patient-to-patient variability, with median within-patient between-region TCR sequence overlap of 62%, 71%, 73%, and 87% for metastatic lesions, and 54% and 77% for tumors (excluding one patient with extremely low sequence numbers) (61). The within-patient TCR sequence overlap between tumors and metastatic sites was 57%, 62%, and 83% (excluding the same patient; ref. 61). Similarly, a study in esophageal cancer found that the average within-patient between-tumor region overlap was 65%, with 26% of clones found throughout the tumor (54). Both studies demonstrated that the degree of overlap between two independent assessments of a single sample is approximately 75%. Thus, the level of overlap observed in these studies is suggestive of a spatially homogeneous, or near homogenous, TILs repertoire. In melanoma, within-patient TCR sequence overlap between lesions ranged from 40% to 80%, but for the top 25 most abundant clones in each patient, clonal abundance was poorly correlated across sites (62).

In lung adenocarcinoma, TCR intratumor heterogeneity was correlated with intratumor neoantigen heterogeneity (59). Similarly, in non–small-cell lung cancer, the number of highly abundant ubiquitous TCRs was correlated with the number of ubiquitous nonsynonymous mutations, while the number of highly abundant regional TCRs was correlated with the number of regional nonsynonymous mutations (53). Both sets of results suggest that TCR spatial heterogeneity may be a consequence of T cells colocalizing with their cognate antigen.

A cross-sample analysis of RNA-seq data from The Cancer Genome Atlas supports this, showing that the number of unique TCR sequences per one thousand unique PCR amplicons in a sample is correlated with the number of nonsynonymous tumor mutations (63). However, in melanoma lesions, no correlation was found between nonsynonymous mutation count and TCR clonal diversity across patients (56) or between nonsynonymous mutation counts and the numbers of unique TCR sequences across lesions within a patient (62).

The above studies did not determine the specificity of the relevant TCR, making it difficult to interpret the results. However, the finding of substantial spatial heterogeneity in the distribution of TCR clones for some cancer types has important implications for the use of lymphocyte-based approaches in the clinic, as elaborated below in IMPLICATIONS.

**Peripheral blood repertoires**

There appears to be little overlap between a patient’s TILs and peripheral blood TCR repertoires (54, 57, 61, 64, 65). Studies quantifying the overlap have found between 19% and 25% overlap, similar to that between healthy tissue and blood (54, 61). One study in esophageal cancer found that even the largest TILs clones are rarely found in blood (54). The TILs–blood overlap may be higher in breast cancer; one study found that approximately 40% of TCRs enriched in TILs relative to healthy breast are found in blood (49), but the higher overlap may be due to the focus on enriched TCRs rather than all TCRs. In addition, although the amount of overlap was correlated with the degree of TILs clonality, some high abundance TILs clones were not found in blood (49). The TIL–blood overlap may also be higher for TCRs ubiquitous in tumors than for those that are regionally restricted, even several months following surgical resection, although the patient-to-patient variability is quite high (53). Finally, one study in patients with melanoma has shown that the CD8⁺ PD-1⁺ subset of peripheral T cells has higher TIL overlap (~50%–60%) in some patients (66). Furthermore, the specific antigens recognized by this subset were similar to that of the TIL populations (66).

Despite the limited overlap between the TILs and peripheral blood repertoires, there is evidence that the peripheral repertoire is altered in patients with cancer and may reflect aspects of the disease. TCR diversity and the number of unique TCR sequences in the blood is lower in patients with cervical cancer than in patients with cervical pre-neoplastic lesions, which is in turn lower than in the blood of healthy women (67). And these values are lower in patients with advanced disease than in patients with early-stage disease (67). This decreased peripheral diversity was attributed to increased clonal expansion, measured by the proportion of sequences accounted for by the top 10 most abundant clones (67). In lung cancer, the diversity of the peripheral T-cell repertoire was lower in patients than in healthy donors, and this held when the analysis was restricted to treatment-naïve patients (68). Peripheral diversity was also inversely correlated with tumor diameter and the number of metastatic sites (68). These results suggest that peripheral T-cell diversity decreases as cancer progresses.

One study in breast cancer found that the peripheral CD8⁺ T-cell repertoire was influenced by the patient’s HER2 expression status (69). Specifically, a larger proportion of highly expanded CD8⁺ clones was found in the blood of patients with HER2⁺ versus HER2⁻ breast cancer (69). In addition, eight sequence motifs in the somatically generated region of the TCR genes were overrepresented in HER2⁺ versus HER2⁻ patients (69). These results are consistent with studies showing HER2 (70) and T cells with HER2 specificity (71) in the blood of patients with breast cancer.
Given the ready accessibility of blood repertoires, it is critical to determine the extent to which they reflect the dynamics of TIL populations and whether they independently provide clinically useful insight into a patient’s antitumor immune response. This is discussed below.

**Sentinel lymph node repertoires**

In both breast and cervical cancer, it has been observed that diversity is lower and clonality is higher in tumor-infiltrating versus sentinel lymph node T-cell populations (58, 60, 67). However, clonal expansion is higher in the lymph nodes of node-positive versus node-negative patients (58). In addition, there is overlap between the two repertoires with more TILs sequences found in tumor-involved than tumor-free sentinel lymph nodes (58, 60).

**Cross-patient sequence sharing**

Identical TCR sequences have been found in TILs populations across multiple patients with the same cancer type (45, 47, 49, 58, 60, 63, 67, 72). Much of the sharing appears to be attributable to public clones, TCR amino acid sequences that are widely shared in the general population, rather than to TCRs responding to a common tumor antigen. This is because many of the sequences were also present in healthy blood samples or have sequence features that tend to mark public clones (e.g., having short sequences with few nucleotides added during somatic generation of the corresponding gene; refs. 49, 60, 63).

**AIRR changes under therapy and associations with clinical outcome**

**Pretreatment AIRR properties associated with outcome**

An urgent need in cancer care is for biomarkers that predict patient responses to therapy. Multiple AIRR profiling studies across multiple cancer types suggest that AIRR profiling may be able to serve as such a biomarker, either alone or in combination with other markers (Table 1). Multiple studies across multiple cancer types have found an association between pretreatment TCR diversity and tumor regression or survival. Two studies in patient cohorts that were heterogeneous with respect to treatment type found that high pretreatment TCR diversity was associated with improved survival (52, 73). The first study considered blood repertoires from breast cancer patients (73); the second study considered adjacent healthy mucosa repertoires from gastric cancer patients (52). In the latter study, a Cox multivariate analysis found that mucosal TCR diversity predicted patient survival independently of all other clinical features (52). High pretreatment TCR diversity in the blood was also found to be associated with better survival in patients receiving anti–CTLA-4 who have hepatocellular cancer (74) or pancreatic cancer (75), as well as in patients with urothelial cancer receiving anti–PD-L1 (76). In patients with hepatocellular cancer, it was also associated with tumor regression (74). Additional studies have found an association between high pretreatment TCR diversity in the blood and tumor regression or longer periods of stable disease in melanoma patients receiving anti–CTLA-4 (77, 78) and in patients with head and neck cancer receiving cetuximab (79). In addition, in patients with cervical cancer under heterogeneous therapies, low numbers of unique TCR sequences in lymph nodes were associated with disease progression (67).

In contrast, response to anti–PD-1 appears to be improved by low pretreatment TCR diversity. Low peripheral TCR diversity was associated with improved survival in melanoma and pancreatic cancer patients receiving anti–PD-1 (75, 78), and in the melanoma patients, low blood or tumor TCR diversity was also associated with tumor regression (78, 80).

Finally, in TILs repertoires from patients with lung adenocarcinoma, high intratumor TCR heterogeneity pretreatment was associated with postsurgical relapse and shorter disease-free survival (59).

**AIRR responses to therapy and their associations with outcome**

To better understand how T-cell responses are modulated in response to therapy, multiple AIRR profiling studies have collected pre- and post-treatment samples to assess changes in AIRR properties under therapy (Table 1; Fig. 1). Therapy tends to induce either repertoire diversification or clonal expansion. Increased TCR diversity in the blood was observed in patients with head and neck cancer receiving cetuximab (79), patients with lung cancer under heterogeneous therapies (68), patients with prostate cancer receiving anti–CTLA-4 and granulocyte/macrophage colony stimulating factor (GM-CSF; ref. 81), patients with melanoma and hepatocellular cancer receiving anti–CTLA-4 (74, 82), and patients with pancreatic cancer receiving either anti–CTLA-4 or anti–PD-1 with or without GVAX (75). In the latter study, the increase in diversity was more pronounced under anti–CTLA-4 than anti–PD-1. In some cases, the repertoire diversification was to the point of large-scale clonal turnover, as in the case of patients with breast cancer on first-line chemotherapy (69), patients with melanoma receiving anti–CTLA-4 (83), and patients with prostate cancer receiving anti–CTLA-4 and GM-CSF (83).

In many of these studies, larger increases in peripheral diversity were associated with a positive therapy response. In ref. 68, larger increases were associated with progression-free survival beyond six months. In ref. 74, higher post-treatment diversity was associated with tumor regression or stable disease with progression-free survival longer than six months. In ref. 79, larger increases in peripheral TCR sequence numbers were associated with tumor regression. In ref. 69, larger divergences from the pretreatment peripheral repertoire were associated with better clinical responses. In ref. 81, a more immediate increase in peripheral TCR diversity was associated with a >50% decline in prostate specific antigen. Unfortunately, this immediate increase was also associated with immune-related adverse events (81), and in ref. 82, larger increases in the number of peripheral TCR sequences and peripheral TCR diversity were associated with treatment-related toxicity.

All of the above studies observing increased TCR diversity in response to therapy were looking in blood. All of the studies looking at the response of TIL populations observed clonal expansion (Table 1; Fig. 1). Clonal expansion of intratumor T-cell populations has been observed in patients with breast cancer receiving neoadjuvant chemotherapy (84), patients with breast cancer receiving anti–CTLA-4 with or without cryoablation (57), patients with melanoma receiving anti–PD-1 (80, 85), and patients with colorectal cancer receiving a cancer peptide vaccine plus oxaliplatin (86). In patients with melanoma receiving BRAF inhibitor therapy, the intratumor clonal expansion was accompanied by the appearance of new clones (87). Finally, a study of patients with urothelial cancer receiving anti–PD-L1 found that, for clones expanding in the tumor that could also be detected in blood, the peripheral clonal abundance also increased under therapy (76).

Large amounts of intratumor clonal expansion under therapy are associated with a positive outcome. In ref. 86, more clonal expansion was associated with longer progression-free survival. In refs. 80, 84, 85, more intratumor clonal expansion was associated with tumor regression or with stable disease with longer progression-free survival. In

---

**Table 1**

<table>
<thead>
<tr>
<th>Cancer Type</th>
<th>Therapy</th>
<th>Clonal Expansion</th>
<th>TCR Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>Anti–PD-1</td>
<td>Increased</td>
<td>Increased</td>
</tr>
<tr>
<td>Lung</td>
<td>Anti–CTLA-4</td>
<td>Increased</td>
<td>Increased</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti–PD-1</td>
<td>Increased</td>
<td>Increased</td>
</tr>
</tbody>
</table>

---

*Published OnlineFirst December 30, 2019; DOI: 10.1158/0008-5472.CAN-19-1457*
<table>
<thead>
<tr>
<th>Cancer type</th>
<th>Treatment</th>
<th>Source tissue</th>
<th>Response to therapy</th>
<th>AIRR property</th>
<th>Outcome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer</td>
<td>Mixed</td>
<td>Blood</td>
<td>—</td>
<td>High pretreatment TCR diversity</td>
<td>Better overall survival</td>
<td>PMC3382902</td>
</tr>
<tr>
<td>Gastric cancer</td>
<td>Mixed</td>
<td>Adjacent healthy mucosa</td>
<td>—</td>
<td>High pretreatment TCR diversity</td>
<td>Better short- and long-term survival</td>
<td>PMC448537</td>
</tr>
<tr>
<td>Lung adenocarcinoma</td>
<td>Mixed</td>
<td>Tumor</td>
<td>—</td>
<td>High pretreatment intratumor TCR heterogeneity</td>
<td>Postsurgical relapse, shorter disease-free survival</td>
<td>PMC5629137</td>
</tr>
<tr>
<td>Cervical cancer</td>
<td>Mixed</td>
<td>Lymph node</td>
<td>—</td>
<td>Low number of unique TCR sequences pretreatment</td>
<td>Disease progression</td>
<td>PMC6262070</td>
</tr>
<tr>
<td>Head and neck cancer</td>
<td>Cetuximab</td>
<td>Blood</td>
<td>—</td>
<td>Increase in the number of TCR sequences</td>
<td>Positive response</td>
<td>PMC6209044</td>
</tr>
<tr>
<td>Melanoma</td>
<td>BRAF inhibitor</td>
<td>Tumor</td>
<td>—</td>
<td>Larger pretreatment number</td>
<td>Positive response</td>
<td>PMC3827093</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>Neoadjuvant chemotherapy</td>
<td>Tumor</td>
<td>—</td>
<td>Clonal expansion, appearance of new clones</td>
<td>Maintenance of pretherapy dominant clones</td>
<td>PMC4922832</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>First-line chemotherapy</td>
<td>Blood</td>
<td>—</td>
<td>Divergence from pretreatment repertoire</td>
<td>Large divergence</td>
<td>PMID:30167801</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>Mixed</td>
<td>Blood</td>
<td>—</td>
<td>Increase in diversity</td>
<td>Progression-free survival &gt; 6 months</td>
<td>PMID:30664810</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Cancer peptide vaccine + oxaliplatin</td>
<td>Tumor</td>
<td>—</td>
<td>Clonal expansion</td>
<td>Longer progression-free survival</td>
<td>PMC4887143</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti–CTLA-4</td>
<td>Blood</td>
<td>—</td>
<td>High pretreatment TCR diversity</td>
<td>Tumor regression or stable disease ≥ 9 months</td>
<td>PMC4469400</td>
</tr>
<tr>
<td>Hepatocellular carcinoma</td>
<td>Anti–CTLA-4</td>
<td>Blood</td>
<td>—</td>
<td>Decrease in clonality/increase in diversity</td>
<td>Better response (RECIST) or stable disease with progression-free survival &gt; 6 months</td>
<td>PMID:50668989</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>Anti–CTLA-4 + cryoablation</td>
<td>Blood</td>
<td>—</td>
<td>Moderate expansion of some clones for the anti–CTLA-4 + cryo arm Anti–CTLA-4 moderate expansion of large number of clones + cryo more large magnitude expansions</td>
<td>Immediate increase in diversity ≥50% decline in PSA; immune-related adverse events</td>
<td>PMC5064859</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>Anti–CTLA-4 + GM-CSF</td>
<td>Blood</td>
<td>—</td>
<td>Increase in diversity</td>
<td>Immediate increase in diversity</td>
<td>PMID:3936199</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti–CTLA-4</td>
<td>Blood</td>
<td>—</td>
<td>Increase in the number of TCR sequences</td>
<td>Toxicity</td>
<td>PMC4008952</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti–CTLA-4</td>
<td>Blood</td>
<td>—</td>
<td>Maintenance of pretherapy dominant clones</td>
<td>Longer overall survival</td>
<td>PMC4588099</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>Anti–CTLA-4 + GM-CSF</td>
<td>Blood</td>
<td>—</td>
<td>Maintenance of pretherapy dominant clones</td>
<td>Longer overall survival</td>
<td>PMC4588099</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti–PD-1</td>
<td>Blood</td>
<td>—</td>
<td>Repertoire turnover</td>
<td>Better response - partial responder (RECIST)</td>
<td>PMID:30325105</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti–CTLA-4</td>
<td>Blood</td>
<td>—</td>
<td>—</td>
<td>Longer progression-free survival</td>
<td>PMID:30325105</td>
</tr>
</tbody>
</table>

(Continued on the following page)
## Table 1. AIRR studies showing associations between outcome and pretreatment AIRR properties or AIRR responses to therapy. (Cont’d)

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>Treatment</th>
<th>Response to therapy</th>
<th>AIRR property</th>
<th>Changes in AIRR profile</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancreatic cancer</td>
<td>Anti-PD-1 + GVAX</td>
<td>High pretreatment clonality (low diversity)</td>
<td>Higher pretreatment clonality</td>
<td>Increase in TCR diversity, more so with anti-PD-1</td>
<td>Longer survival (6 months)</td>
</tr>
<tr>
<td></td>
<td>Anti-CTLA-4 + GVAX</td>
<td>High pretreatment clonality (low diversity)</td>
<td>Higher pretreatment clonality</td>
<td>Increase in TCR diversity, more so with anti-CTLA-4</td>
<td>Longer survival (6 months)</td>
</tr>
<tr>
<td></td>
<td>Anti-PD-1</td>
<td>High pretreatment clonality (low diversity)</td>
<td>Higher pretreatment clonality</td>
<td>Increase in TCR diversity, more so with anti-PD-1</td>
<td>Longer survival (6 months)</td>
</tr>
<tr>
<td>Urothelial cancer</td>
<td>Anti-PD-L1</td>
<td>High pretreatment clonality (low diversity)</td>
<td>Higher pretreatment clonality</td>
<td>Increase in TCR diversity, more so with anti-PD-L1</td>
<td>Longer survival (6 months)</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Anti-PD-1</td>
<td>High pretreatment clonality (low diversity)</td>
<td>Higher pretreatment clonality</td>
<td>Increase in TCR diversity, more so with anti-PD-1</td>
<td>Longer survival (6 months)</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>Anti-PD-1 and GM-CSF</td>
<td>High pretreatment clonality (low diversity)</td>
<td>Higher pretreatment clonality</td>
<td>Increase in TCR diversity, more so with anti-PD-1</td>
<td>Longer survival (6 months)</td>
</tr>
</tbody>
</table>

Note: AIRR properties shown in bold are those that are associated with outcome pretreatment.

ref. 76, substantial expansion was associated with progression-free survival longer than six months. In addition, in patients with pancreatic cancer receiving checkpoint blockade therapy with or without GVAX, high post-treatment clonality in the periphery under anti-PD-1 and a large number of expanded clones in the periphery post-treatment under anti-CTLA-4 are both associated with survival beyond six months (75).

Several studies suggest that, whether there is repertoire diversification or clonal expansion in response to therapy, the presence of tumor-specific clones before therapy may be important to having a positive response to therapy. In patients with melanoma receiving BRAF inhibitor therapy, T-cell clonality in the tumor increased under therapy, and expansion was primarily among clones not observed before therapy, suggesting significant, therapy-induced recruitment of new clones into the tumor (87). However, it was the maintenance of clones that were dominant before therapy that was associated with a positive treatment response (87). In patients with lung cancer receiving various therapies, peripheral T-cell diversity increased under therapy, and the increase was associated with progression-free survival beyond six months, as described above. However, a high overlap between the pre- and post-treatment peripheral repertoires was independently associated with progression-free survival beyond six months (68). Finally, in patients with melanoma receiving anti–CTLA-4 and patients with prostate cancer receiving anti–CTLA-4 plus GM-CSF, the maintenance of clones that were dominant pre-therapy was associated with longer overall survival (83). Although the specificity of the pretreatment clones was not determined, these results all suggest that the positive therapy response may be due to preexisting tumor-specific T-cell clones.

The implications and potential utility of AIRR associations with treatment response and outcome are discussed below.

### The diagnostic potential of AIRR

Methods for cancer early detection are urgently needed, particularly methods that can be applied to blood. Recently, Ostmeyer and colleagues proposed an approach in which cancer-specific AIR would be assayed in blood using AIRR deep sequencing. To assess the feasibility of this approach, they determined whether TILs repertoires are distinguishable from repertoires associated with healthy tissue of the same type based on TCR sequence patterns alone. Their approach achieved classification accuracies of 93% and 94% for colorectal cancer and breast cancer, respectively, including sensitivities of 100% in both cases (88). The approach relies on finding TCR with biophysicochemical motifs that are shared across TILs repertoires from patients with the same cancer type but absent in the TCR repertoires from patient-matched adjacent healthy tissue. Importantly, the colorectal cancer and breast cancer motifs are different and specific to the cancer type. It is not yet known whether the method can be applied to blood repertoires, however, the results are promising and suggest that AIRR deep sequencing may represent an important new resource for the development of cancer early detection assays.

### Implications and Open Questions

The above observations suggest that T cells traffic through healthy tissue into tumors, where those with specificity for tumor antigens undergo clonal expansion. These expanding T-cell clones, under continued exposure to antigen, become exhausted. In response to therapy, including therapies that do not directly target T cells, the TILs population tends to undergo further clonal expansion while the peripheral population tends to undergo diversification...
Patients with a measurable response in either population tend to have a positive response to therapy, as do patients who maintain clones that were present before therapy (Table 1). For treatments other than anti–PD-1, high levels of pretreatment TCR diversity appear to be associated with positive responses to therapy. This may reflect the fact that patients with high diversity have a broader spectrum of antigen specificities to draw from and are therefore more likely to have pre-existing clones with tumor specificity. This is supported by studies indicating improved outcomes for patients who maintain pretreatment clones in their post-treatment repertoire. If, however, as suggested above, TCR diversity decreases as cancer progresses, high pretreatment diversity could simply be a marker of early disease, and early-stage patients can be expected to have better outcomes. Thus, studies aimed at disentangling the effects of pretreatment TCR diversity, pre-existing tumor-specific clones, and disease stage are needed.

In contrast, low pretreatment diversity is associated with a positive outcome for anti–PD-1 therapy. This may be related to the finding that the CD8+ PD-1+ TILs subset is highly expanded and includes many clones with tumor reactivity. Thus, patients with low diversity, and simultaneously higher clonality, may have repertoires dominated by tumor-specific CD8+ PD-1+ cells that can be reinvigorated under anti–PD-1 therapy. This population of cells may not be reinvigorated, or as readily reinvigorated, under other therapies. Again, studies aimed at determining the precise effects are needed.

These results suggest that assessing pretreatment T-cell diversity and clonality, and whether clonality is concentrated within the PD-1+ subset, could help guide therapy decisions for some patients. However, these results also suggest that strategies for identifying patients capable of undergoing diversification or expansion are needed. Additionally, determining pretreatment whether patients have tumor-specific clones may be critical. The answers to these questions could form a systematic strategy for determining which therapy or combination of therapies is most likely to elicit a robust antitumor T-cell response from a patient’s in vivo T-cell population. The answers could also inform which patients may need ACT, whether ACT from their own cells will likely succeed, or whether engineered TCRs may need to be leveraged.

The finding of substantial spatial heterogeneity in the distribution of clones throughout tumors and metastatic sites for at least some cancer types has important implications for the use of lymphocyte-based approaches in the clinic. Profiling the location, density, and subset composition of TILs is widely accepted to be of prognostic value for some cancer types (23). Evidence described here suggests that TIL diversity and clonality may also be of prognostic value. Careful studies are needed to determine whether multi-region approaches can improve the prognostic accuracy of lymphocyte measures or whether they are in fact required in order to achieve reliable results for some cancer types. Indeed, cancer types for which negative results have been obtained should perhaps be revisited with a multi-region approach. Similarly, studies collecting TILs for autologous ACT therapy might benefit from the use of material from distributed tumor regions and/or metastatic sites.

Given the accessibility of blood relative to tumor tissue, it is important to determine which features of antitumor immune responses can be assayed from blood, particularly for post-treatment monitoring where serial sample collection may be needed. The answer may depend on the cancer type and the extent to which an individual tumor has been vascularized. Results so far indicate that the peripheral T-cell repertoire may become less diverse as cancer progresses and that the level of peripheral diversity may be prognostic. They further indicate that tumor-specific T cells are rare in blood (89), with a possible exception being cases where a tumor-associated antigen, such as overexpressed HER2, can be found in high levels in the blood. Tumor-specific clones found in the blood may be concentrated in the CD8+ PD-1+ subset (66). Thus, studies elucidating the dynamics of this peripheral subset are needed.

Methods for cancer early detection are urgently needed, particularly methods that can be applied to blood. Most efforts have focused on approaches to directly detect cancer (e.g., detecting circulating tumor cells or cancer mutations in circulating cell-free DNA) or cancer products (e.g., CA-125 or prostate specific antigen). Although the results are promising in some cases, particularly when combining multiple markers, the approaches generally struggle with low sensitivity, particularly when looking at early stage disease (90). The recent article by Ostmeyer and colleagues (88, 91), although focused on tissue repertoires, demonstrates that AIRR sequencing may have an important role to play here too.

Caution in interpreting the above results is needed, particularly when comparing results across studies. Most AIRR sequencing studies conducted so far in the cancer domain have used relatively small sample sizes and heterogeneous patient populations. Oftentimes, the populations are poorly described with respect to things such as therapies, sampling time points, and response measures. Interpretation is further exacerbated by high sample-to-sample variability in many cases. In addition, AIRR deep sequencing is a relatively young technology; the sequencing and analysis protocols are still highly variable and evolving rapidly. One critical point for consideration is whether studies sequenced RNA or genomic DNA. Conclusions that rely on interpretations of clonal abundances are expected to be unreliable in RNA-based studies, because there is no correspondence between the number of sequencing templates and the number of cells. However, many RNA-based AIRR studies report measures, such as diversity and clonality, that rely for their own accuracy on reasonably accurate estimates of clonal relative abundances. While in principle a correspondence between cell and template abundance is preserved in studies based on genomic DNA, it is confounded by differences in amplification efficiencies between different primer sets, as well as between primers within a single primer set. Thus, for studies for which clonal relative abundances are important measures, it is highly recommended to use sequencing protocols with built-in strategies that allow reliable estimates of clonal abundance (92–94). There is currently tremendous variability between analysis protocols, as well, including in the most basic data processing steps (e.g., which regions of AIRR sequence are used to infer clonal membership). The downstream analyses can also differ substantially, even in the calculation of widely used measures. For example, the term “diversity” is used in some studies to refer to the number of unique TCR sequences observed, while this is properly referred to as richness in some studies. In some studies, diversity is estimated using formal diversity indices, such as the Shannon diversity index, but different indices are used in different studies. In addition, they are often improperly deployed. They rely on accurate estimates of clonal relative abundance but are sometimes used in RNA-based studies or genomic DNA-based studies where primer amplification bias was not properly controlled for. Similarly, clonality is sometimes estimated by counting the number of highly expanded clones, with variable definitions for what constitutes being highly expanded. Clonality may also be estimated by the proportion of total sequences accounted for by some number of the largest clones (e.g., top 10) or by subtracting a normalized Shannon’s diversity index from one, as described above. Finally, when the percent overlap between two repertoires is calculated, there is inconsistency
regarding whether the denominator is the total number of sequences in one of the repertoires or in both repertoires summed.

Given the high variability in sequencing and analysis protocols, it is important to understand differences in the underlying approaches before concluding that any two studies reached similar or different conclusions.

**Future Directions**

The most important questions in cancer research include (i) improving our understanding of why natural, antitumor immunity fails; (ii) developing new approaches for potentiating natural, antitumor lymphocyte responses; (iii) improving our understanding of how lymphocyte populations are modulated by therapeutic intervention, along with our ability to assess their response potential to improve the matching of patients to therapies; (iv) improving our ability to design personalized treatments; and (v) developing streamlined strategies for implementing them. The results presented above show that AIRR deep sequencing will be a critical component to advancing these research goals. To do so, however, the field needs to move towards studies using larger numbers of patients, as well as methods that better support cross-study comparisons, comprehensive profiling strategies that profile lymphocyte populations along multiple axes, and more sophisticated analysis approaches.

To better facilitate the joint interpretation of results across multiple studies, it is highly recommended that any studies with an AIRR sequencing component adhere to standards being developed by the AIRR Community (https://www.antibodysociety.org/the-airr-community/), a community of AIRR researchers formed in recent years with the primary purpose of developing standards that will enhance the reproducibility, comparability, and reusability of AIRR sequencing data (95–98). Additional articles describing Community standards are forthcoming.

Multi-axis profiling of lymphocyte populations in cancer is urgently needed. To fully address the questions outlined above, we need approaches that allow simultaneous assessment of each cell’s functionality (e.g., regulatory versus cytotoxic T cell) and functional potential (e.g., activated versus exhausted), along with both chains of its AIR heterodimer and the corresponding antigen specificity. Recent technological advancements are making such comprehensive profiling feasible in the near term. For example, single cell RNA-seq technologies allow the capture of cellular phenotypic information together with both AIR chains (99, 100). Although this represents a significant benefit over bulk sequencing approaches, there are additional benefits, including improved correction of sequencing error and more robust estimates of clonal abundance. The current limitation is in throughput, which, although it has increased in recent years, is still low relative to bulk sequencing approaches. However, increasing throughput is an active area of research, and the throughput of single-cell approaches is expected to be competitive in the future (101–103).

Advances in identifying tumor antigens and TCRs with tumor antigen specificity have also been made. The current most widely used approach for identifying tumor antigens is to apply bioinformatics pipelines to whole exome or whole transcriptome data to identify tumor-specific, nonsynonymous mutations or genes over-expressed in tumor cells, followed by prediction of peptides expected to be presented by MHC I molecules (104, 105). These predictions should be followed by in vitro validation to determine which of the predicted
peptides is actually presented by a given MHC I allele. Mass spectrometry of peptides eluted from MHC I molecules allows for a more comprehensive identification of presented peptides (106). In both cases, however, the immunogenecity of the presented peptides is not known. Although T-cell functional assays can be combined with TCR sequencing to identify immunogenic peptides and the TCRs that bind them, there have been no assays that enabled this at scale. Technological advances based on single cell RNA-seq are addressing this problem as well. For example, a relatively new approach called TetTCR-seq will allow multimer sorting of T cells with as many as approximately 1,000 peptides that are DNA barcoded, allowing high-throughput peptide processing followed by sequencing to match peptides with the sequences of bound TCRs (107).

These enhanced capabilities will be especially important for improving the applicability and effectiveness of personalized treatments, particularly ACT with engineered T cells. Although ACT with autologous TILs has resulted in durable, complete responses in some patients with metastatic melanoma (26, 108), extending this success to all melanoma patients or to other cancer types has proven challenging. The key challenges include the rarity of TILs expressing TCR with high affinity for tumor-specific antigens and the failure of transferred cells to expand in vivo, accumulate in tumors, and maintain their cytotoxicity. Synthetic biology is offering exciting ways to overcome these obstacles by engineering cells with particular phenotypes and antigen specificities (109). T cells can be engineered to express a specific TCR that was previously identified has high-affinity and tumor-specific in a patient successfully treated with ACT or that was derived from immunizing mice with patient tumor (28, 110, 111). In addition, T cells can be engineered to express chimeric antigen receptors (CAR), which combine part of the TCR signaling complex as the intracellular domain, a transmembrane domain, and a single-chain variable fragment that encodes a single antibody light chain variable region and a single antibody heavy chain variable region as the extracellular antigen-binding domain (112, 113). CAR T cells have been quite successful for treating B cell malignancies (112). They have been less successful against solid tumors but offer hope of targeting surface molecules, particularly in cases where cancer immune evasion has resulted in downregulated expression of MHC I (112). AIRR sequencing and the newer multi-axis technologies should facilitate the identification of AIRR genes that would be good candidates for engineered receptors.

Finally, the overwhelming majority of AIRR sequencing studies are primarily descriptive and apply only the most basic analysis techniques of calculating repertoire-level statistics, such as diversity and clonality, and looking for sequence sharing across individuals by directly comparing amino acid sequences. To penetrate and utilize the vast amounts of information contained in AIRR, and to integrate it with that stemming from the new, multi-axis profiling strategies, new analysis methods are needed. Several research groups have been actively developing new approaches that leverage both sophisticated learning algorithms and innovative ways to represent AIRR (88, 114–119). Further advances are needed in this area.

In summary, AIRR studies have contributed a foundation of knowledge about the lymphocyte population dynamics in cancer and have pointed the way to important, unanswered questions. Further studies are needed to determine which of the findings will replicate and which will generalize across cancer types versus being specific to one or a few cancer types. A new era of large studies using multi-axis lymphocyte profiling and innovative analysis methods has the potential to make significant contributions toward answering the current most pressing questions in cancer research.

Disclosures of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Received May 10, 2019; revised October 14, 2019; accepted December 17, 2019; published first December 30, 2019.

References

Adaptive Immune Receptor Repertoire Profiling in Cancer


The Diagnostic, Prognostic, and Therapeutic Potential of Adaptive Immune Receptor Repertoire Profiling in Cancer

Lindsay G. Cowell


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-19-1457

Cited articles
This article cites 116 articles, 26 of which you can access for free at:
http://cancerres.aacrjournals.org/content/80/4/643.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerres.aacrjournals.org/content/80/4/643.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.