Sarcoma background & epidemiology

Sarcomas are thought to originate from mesenchymal tissue, and account for 1% of adult solid malignancies and over 20% of pediatric solid malignancies. Nearly 11,000 people in the USA and 200,000 people in the world will be diagnosed with sarcomas each year [101]. When they metastasize, sarcomas typically have a very poor outcome. Unlike other types of tumors, the molecular basis and accompanying genomic changes in sarcoma are not well understood. This is due to the rarity of sarcoma and the difficulty in collecting enough fresh specimens for genomic analysis, which remains especially true for metastatic and relapsed lesions. Recurrent sarcomas have evolved the capacity to spread and resist therapy, although we still know very little about what drives their aggressiveness. As next-generation sequencing and other genome-wide technology becomes more readily available and affordable, we will begin to learn more about sarcomas and, thereby, increase our ability to identify candidate genes and pathways for therapeutic targeting.

The molecular mechanisms for sarcomagenesis can be divided into three broad categories, as described by Taylor et al. in a recent review of sarcoma genomics and therapeutic targets [1]. These categories include: translocations that create fusion proteins leading to transcriptional dysregulation; somatic mutations; and DNA copy number changes. Advances in molecular technology have defined the genome-wide targets and transcriptional effects of sarcoma fusion proteins, starting initially with gene expression microarrays and more recently with chromatin immunoprecipitation arrays and sequencing. Sarcomas can be categorized in several different ways. When creating a taxonomy for classification, soft tissue sarcomas can be divided by tissue of origin, by prognosis or even by specific driver alterations [1]. However, more typically, sarcomas are divided into two categories: balanced translocation-associated sarcoma (BTAS; genome stable); and complex genotype–karyotype sarcoma (genome unstable) [2]. Each BTAS has its own recurring translocation. For instance, 85% of Ewing sarcomas will contain the t(11;22) (q24;q12) EWSR1/FLI1 translocation, and the remaining tumors will have EWSR1 bound to a different ETS protein (e.g., ERG, ETV1, ETV4 or FEV). Other BTASs include clear-cell sarcoma, desmoplastic small-round-cell tumor, myxoid chondrosarcoma, myxoid liposarcoma, alveolar rhabdomyosarcoma (ARMS), synovial sarcoma, dermatofibrosarcoma protuberans, congenital fibrosarcoma, inflammatory myofibroblastic tumor and alveolar soft-part sarcoma. The complex genotype–karyotype sarcoma category includes tumors with multiple genomic alterations such as osteosarcoma, undifferentiated pleomorphic sarcoma (formerly called malignant fibrous histiocytoma), embryonal rhabdomyosarcoma, leiomyosarcoma, malignant peripheral nerve sheath tumor, angiosarcoma, fibrosarcoma (other than congenital), chondrosarcoma (other than extraskeletal myxoid) and liposarcoma (other than myxoid). Helman and Meltzer have written a very thorough and clear review on the mechanisms of sarcoma development, including an in-depth discussion of the molecular basis of BTAS and complex genotype–karyotype sarcoma tumors [2].

Genomic landscape of sarcomas

The genomic landscape of somatic copy number alterations (CNAs) across human cancers...
is now being explored by both microarray and next-generation sequencing platforms. Recently, Beroukhim et al. categorized at least 26 different tumors based on genomic CNAs, including several sarcomas [3]. Interestingly, all of the cancers clustered into four categories based on hematopoietic, neural, sarcomas and epithelial lineage. The sarcomas clustered into two different subcategories: synovial sarcoma, mesothelioma and gastrointestinal stromal tumor (GIST); and leiomyosarcoma, pleomorphic liposarcoma, undifferentiated pleomorphic sarcoma and dedifferentiated liposarcoma.

“As next-generation sequencing and other genome-wide technology becomes more readily available and affordable, we will begin to learn more about sarcomas and, thereby, increase our ability to identify candidate genes and pathways for therapeutic targeting.”

Barretina et al. further explored genomic alterations in sarcoma by performing an integrative analysis of targeted sequencing, CNAs, and mRNA expression (n = 207, from seven sarcoma subtypes) [4]. Using this approach, they discovered recurrently altered genes that could hold therapeutic promise for subtype-specific therapeutic targets. These included TP53 mutations (17% pleomorphic liposarcomas), NF1 mutations (10.5% myxofibrosarcomas and 8% pleomorphic liposarcomas) and PIK3CA mutations (18% myxoid/round-cell liposarcomas). As more sarcomas are collected for analysis, more of these recurring targets can be found and incorporated into preclinical and clinical trials. If we are fortunate, then targets may be discovered for which existing drugs already exist, shortening the time before patients can receive these therapies.

Rhabdomyosarcoma translocations

Even with multiple samples and known genomic rearrangements, the observed significance is not always clear. For example, many investigators have explored the association with clinical outcome for the recurring PAX3–FOXO1(FKHR) and PAX7–FOXO1(FKHR) translocations in ARMS, with varying results. In a study of 78 patients from the Intergroup Rhabdomyosarcoma Study (IRS)-IV/Children’s Oncology Group (COG) trial, Sorensen et al. found a clear pattern of worse prognosis for ARMS patients with PAX3–FOXO1 gene fusions [5]. This difference was particularly pronounced for those with metastatic disease, although the numbers were small. Several years later, Stegmaier et al. reported on the unclear significance of PAX3–FOXO1 status when they found better outcome of this translocation in localized ARMS compared with PAX7- or translocation-negative patients and no clinical difference in metastatic disease depending on PAX3 or PAX7 status [6]. This study included 121 patients from the Cooperative Soft Tissue Sarcoma Group (CWS). More recently, Missiaglia et al. reported that PAX3–FOXO1 is indeed the key prognostic molecular marker and improved risk stratification in ARMS in a large study based on 287 patients from both the European Paediatric Soft Tissue Sarcoma Study Group (EpSSG) and COG [7]. They also described an improved clinico- and molecular risk score that incorporates translocation status, IRS Tumor Node Metastasis (TNM) stage and age [7]. However, the debate is far from settled and recent articles warn about the universal interpretation of these reports due to different cohorts studied with different therapeutic strategies, including selection bias of available samples for molecular analysis within cohorts [8–10]. One way to answer this seemingly straightforward but actually quite challenging question would be to perform a prospective molecular analysis on every patient enrolled in a large, multi-institutional, international cooperative trial.

SDH-deficient GISTs

A recent genomic discovery in the field of sarcoma research is the absence of mitochondrial complex II (SDH) in nearly 100% of wild-type (pediatric) GISTs. Wild-type GIST, defined as lack of somatic KIT or PDGFRA mutations, is an exceedingly rare sarcoma. It most often occurs in children and adolescents, but even so, its true incidence in the population is still unknown [11]. In response to this, the NIH established the collaborative Pediatric and Wild-type GIST Clinic in 2008 to study the clinical and molecular features of this very rare sarcoma [102]. Initially held biannually, this clinic offers a multidisciplinary approach that includes oncologists, surgeons, endocrinologists, radiologists, geneticists and genetic counselors. Importantly, this NIH clinic includes blood and tumor collection, and over 70 participants. Taking advantage of this collection of rare sarcomas, Janeway et al. demonstrated that the loss of SDHB protein expression (part of the SDH complex) is found in 100% of wild-type GIST patients [12]. This contrasts with KIT-mutant GISTs where 67% of patients
showed normal SDHB expression. Others have since validated these findings and, like Janeway et al., have demonstrated SDHB protein loss in wild-type GISTs is associated with SDH germline mutations [13,14]. This has led to a recent name change of wild-type GISTs to ‘SDH-deficient GISTS.’ This genomic discovery has clinical implications as patients with SDH-deficient GISTS are at high risk for familial paraganglioma and pheochromocytoma syndrome due to inherited SDH mutations. These patients and their affected family members can now be entered in tumor surveillance programs that can detect early tumors prior to symptoms [15].

**Genome-wide copy number & outcome**

Perhaps the best demonstration of the clinical translation of genome-wide data in sarcoma is the repeated finding of increased genomic instability associated with worse event-free and overall survival. This was recently shown to be true in a small cohort of 40 patients from Utah (USA) with Ewing sarcoma [16]. Using novel single nucleotide polymorphism microarray technology, Jahromi et al. demonstrated that patients with at least one copy number change in any of eight specific regions had an overall survival of 39% compared with 100% in those patients lacking any of these eight CNAs (p < 0.001) [16]. These samples were obtained from clinically archived formalin-fixed paraffin-embedded tumor blocks with over 12 years of follow-up. Similar trends of worse survival with increasing copy number have been reported for osteosarcoma [17], malignant peripheral nerve sheath tumor [18,19] and even across multiple sarcoma subtypes [20]. Despite their success in individual cohorts, these CNA prognostic signatures have yet to be tested in prospective clinical trials and used in current studies to risk-stratify patients. This is most likely due to the continued problem of the rarity of sarcomas and difficulty obtaining enough samples for validation.

In summary, genomic discovery is rapidly expanding in both sarcoma and other cancer types. The field of sarcoma remains challenged by the limited number of samples available for both discovery and validation. Genome-wide tools offer the very real promise of discovery of new mechanisms for sarcomagenesis, identification of novel candidate pathways for therapeutic targeting and even preventative strategies for those patients most at risk of developing tumors. In order to realize this promise, sarcoma researchers and clinicians can learn from the examples highlighted above. Understanding the genomic landscape of sarcomas will help to identify candidate regions for drugs and sarcoma research (copy number studies). In addition, patient selection for genomic research is critical and sometimes controversial (e.g., rhabdomyosarcoma translocations and outcome). We need to work together to establish comprehensive cohorts of rare tumors to study (e.g., SDH-deficient GIST Clinic), and all results need to be replicated and validated (i.e., using prognostic gene signatures). These collaborations will be key for moving the field forward as we work together to improve the outcome of our patients with sarcoma.

**Financial & competing interests disclosure**

JD Schiffman is a Damon Runyon Clinical Investigator supported (in part) by the Damon Runyon Cancer Research Foundation (CI#56-11). He has received speaking honoraria from Affymetrix, Inc. (Santa Clara, CA, USA). The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

**References**


Websites