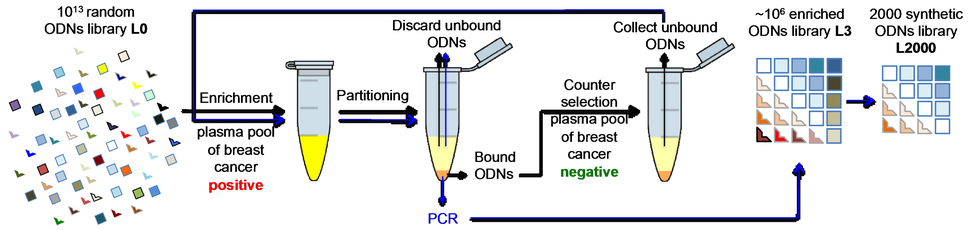
Adaptive dynamic artificial poly-ligand targeting (ADAPT) enables plasma-based exosome profiling with potential diagnostic utility

Valeriy Domenyuk[1], Zhenyu Zhong[1], Adam Stark[1], Nianqing Xiao[1], Heather O'Neill[1], Jie Wang[1], Xixi Wei[1]; Teresa Tinder[1], Janet Duncan[1], Andrew Hunter[1], Mark R. Miglarese[1], Joachim Schorr[1], David Halbert[1], John Quackenbush[2], George Poste[3], Günter Mayer[1,4], Michael Famulok[1,4,5], and David Spetzler[1].

[1]Caris Life Sciences, Phoenix, AZ, USA; [2]Department of Biostatistics and Computational Biology, Dana-Faber Cancer Institute and Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA; [3]Complex Adaptive Systems Initiative, Arizona State University, Scottsdale, AZ, USA; [4]LIMES Program Unit Chemical Biology & Medicinal Chemistry, University of Bonn, Germany; [5]Chemical Biology Max-Planck-Fellowship Group, Center of Advanced European Studies and Research (CAESAR), Bonn, Germany.

Introduction: Single stranded DNA (ssDNA) libraries consisting of several trillion oligodeoxynucleotides (ODNs) can adopt a nearly infinite number of three-dimensional structures. These structures can potentially bind any biomolecule and can be screened for specificity toward important biomarkers by employing suitable enrichment schemes. Since no prior knowledge on the binding partner is required, massively parallel biomarker identification is possible even on complex matrices like biological fluids and across a wide range of biological conditions. Here we present Adaptive Dynamic Artificial Poly-ligand Targeting (ADAPT) as a platform for biomarker and target discovery. We employed ADAPT for the molecular profiling of exosome-associated proteins in small volume plasma samples from women with breast cancer and healthy donors.

Enrichment of aptamer library for ADAPT



ADAPT workflow

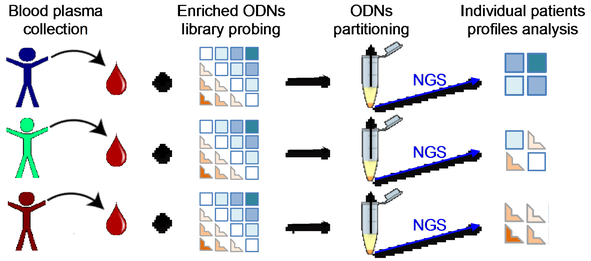


Figure 1. ADAPT principle: Random ssDNA-libraries (L0) of 10^{13} unique ODNs were subjected to a number of selection and counter-selection steps on pooled blood plasma of breast cancer and healthy women. Several positive and negative enrichment schemes were employed, and exosome isolation and ODNs library partitioning were performed by ultracentrifugation and/or PEG precipitation. After library enrichment reduction of complexity to 10^6, ODN library (L3) was used to probe an independent set of individual plasma samples from women with or without breast cancer. 2000 differentially-binding ODNs with significant p-values were re-synthesized and combined in equimolar amounts to create a profiling library (L2000). The L2000 library was used to probe plasma samples from 323 individuals (206 from breast cancer patients and 117 from healthy donors) in triplicate. Next Generation Sequencing (NGS) was used to quantitated bound ODN from each plasma sample.

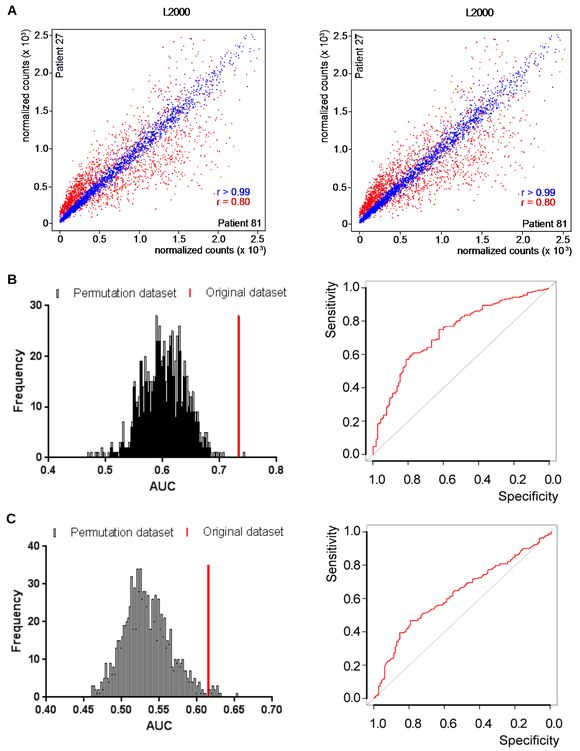


Figure 2. ADAPT characterization and evaluation on 323 clinical samples: **(A)** Distribution of normalized counts of aptamers recovered from ADAPT enriched library L3 and L2000 on technical replicates from the same sample (blue dots) compared to average counts from 3 replicates of two non-related samples (red dots). **(B)** Random-Forest (RF) Out-of-Bag (OOB) ROC AUC from 323 clinical samples and permutation analysis of its reliability; the ROC AUC in the original dataset is 0.73, which is significantly higher compared to the majority of 1000 permutations (p=0.001). **(C)** A more strict 10-fold Cross-Validation ROC AUC in the original dataset is 0.63, which is significantly higher compared to the majority of 1000 permutations (p=0.014).

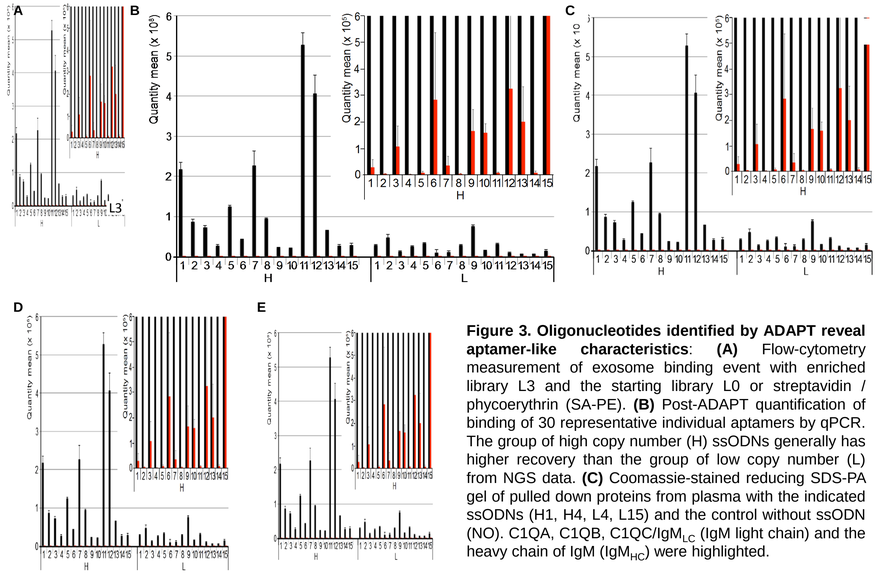


Figure 3. Oligonucleotides identified by ADAPT reveal aptamer-like characteristics: **(A)** Flow-cytometry measurement of exosome binding event with enriched library L3 and the starting library L0 or streptavidin / phycoerythrin (SA-PE). **(B)** Post-ADAPT quantification of binding of 30 representative individual aptamers by qPCR. The group of high copy number (H) ssODNs generally has higher recovery than the group of low copy number (L) from NGS data. **(C)** Coomassie-stained reducing SDS-PA gel of pulled down proteins from plasma with the indicated ssODNs (H1, H4, L4, L15) and the control without ssODN (NO). C1QA, C1QB, C1QC/IgM$_{LC}$ (IgM light chain) and the heavy chain of IgM (IgM$_{HC}$) were highlighted.

(D) Filter retention analysis of C1Q-binding by the ssODNs H1 and H11 at indicated C1Q concentrations. H11RC, the reverse complement of H11, is used as control. **(E)** ELONA analysis of C1Q-binding by the ssODNs H11 at indicated aptamer concentrations and fixed C1Q concentration (0.625 nM). H11RC as control shows low binding and no aptamer control shows background of detector Streptavidin-HRP. H11 specifically binds C1Q (KD = 44.4 nM).

Summary:

- ✓ We have demonstrated the feasibility of aptamer library enrichment directly on blood plasma and have identified a set of 2000 DNA aptamers that distinguish plasma from women with breast cancer from women without breast cancer.
- ✓ This liquid biopsy approach requires only 200 microliters of plasma and is amenable to high-throughput processing.
- ✓ By employing a number of statistical approaches including rigorous cross-validation, we consistently achieve ROC AUC values >0.6.
- ✓ ADAPT – derived breast cancer test may serve as a vital diagnostic adjunct that can be easily incorporated into standard clinical practice.