HHS Public Access

Author manuscript

ACS Nano. Author manuscript; available in PMC 2022 January 08.

Published in final edited form as:

ACS Nano. 2021 March 23; 15(3): 5631–5638. doi:10.1021/acsnano.1c00782.

Sequencing-Based Protein Analysis of Single Extracellular Vesicles

Jina Ko[1,2,#], **Yongcheng Wang**[2,3,4,#], **Kuanwei Sheng**[2], **David A. Weitz**[2,3], **Ralph Weissleder**[1,5,*]

[1]Center for Systems Biology, Massachusetts General Hospital, 185 Cambridge St, CPZN 5206, Boston, MA 02114, USA

[2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA

[3]John A. Paulson School of Engineering and Applied Sciences and Department of Physics, Harvard University, Cambridge, MA 02138, USA

[4]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

[5]Department of Systems Biology, Harvard Medical School, 200 Longwood Ave, Boston, MA 02115, USA

Abstract

Circulating extracellular vesicles (EV) - biological nanomaterials shed from most mammalian cells - have emerged as promising biomarkers, drug delivery vesicles, and treatment modulators. While different types of vesicles are being explored for these applications, it is becoming clear that human EV are quite heterogenous even in homogenous or monoclonal cell populations. Since it is the surface EV protein composition that will largely dictate their biological behavior, high throughput single EV profiling methods are needed to better define EV subpopulations. Here we present an antibody-based immuno sequencing method that allows multiplexed measurement of protein molecules from individual nanometer sized EV. We use droplet microfluidics to compartmentalize and barcode individual EV. The barcodes/antibody-DNA are then sequenced to determine protein composition. Using this highly sensitive technology, we detected specific proteins at the single EV level. We expect that this technology can be further adapted for multiplexed protein analysis of any nanoparticle.

Keywords

sequencing; extracellular vesicles; droplet microfluidics; high throughput; multiplexing

[*]R. Weissleder, MD, PhD, Center for Systems Biology, Massachusetts General Hospital Research Institute, 185 Cambridge St, CPZN 5206, Boston, MA, 02114, 617-726-8226, rweissleder@mgh.harvard.edu.

[#]These authors contributed equally to this work.

SUPPORTING INFORMATION AVAILABLE: <Bead-DNA$_{BC}$ and Ab-DNA$_{BC}$ sequences. Microfluidics chip design. Control experiment antibody specificity.> This material is available free of charge *via* the Internet at http://pubs.acs.org.

INTRODUCTION

Circulating extracellular vesicles (EV) are typically <1,000 nm in size, occur at concentrations of up to 10^{7-11} vesicles/ml of peripheral blood in patients, are fairly stable over time,[1] and have been shown to contain small amounts of proteins and nucleic acids reflective of those found in parental cells.[2,3] The vesicles differ in size, molecular composition, biogenesis, and function.[4,5] EV include exosomes and microvesicles among other membrane vesicles.[6–8] EV are not only shed by tumor cells (tEV) but also by host cells (hEV). Furthermore, bulk EV protein content has been shown to vary temporally and recent studies have shed light on the composition of individual vesicles investigating mostly abundant proteins.[9–11] One emerging view is that the protein expression in well defined vesicle populations (*e.g.* exosomes only) varies considerably from one vesicle to the next. Given this stochastic biomarker expression and scarcity of certain proteins in vesicles, highly sensitive methods of single EV analyses are needed.

A number of different analytical methods have been developed to analyze EV,[12–14] most of them relying on bulk measurements requiring ~$10^{3–6}$ EV for analysis. Yet, the identification of a small number of tumor originating vesicles (such as those found in early cancers) in a background of host EV may be impossible by bulk methods. One way to solve the problem is to develop single (“digital”) EV analysis techniques. Such single EV analysis could be extremely valuable not only for early detection but also for studying tumor heterogeneity and phenotypic changes occurring during therapy. Because of the unmet need for single vesicle analysis, there has been increasing interest in this challenge. Some recent approaches of single vesicle analyses have included optical trapping,[15] Raman spectroscopy,[16] flow cytometry,[17,18] and cyclic imaging.[10] So far, the latter method allows rapid multiplexed protein analysis in individual vesicles. However, optical sensing alone has limitations such as limited amplification (sensitivity), limited multiplexing, and perhaps a lower throughput.

Here, we overcome the sensitivity limitation and increase multiplexing and throughput by using a sequencing-based single EV protein profiling method. The approach borrows from single cell RNA sequencing (scRNAseq) and which has been highly successful in analyzing whole cells.[19–21] In contradistinction to scRNAseq however, we faced a number of challenges: i) an average exosome has a ~10^6 times smaller mass compared to a single cell, ii) our primary interest was in protein profiles rather than endogenous mRNA since the latter can be rare in single EV[22] and it is the protein composition that defines pharmacological and physiological behaviors, iii) that the actual number of different proteins in individual EV is exceedingly low, and iv) there are no good accepted gold standards to compare measurements against. We were further interested in developing a method that would allow one to profile thousands of EVs and potentially dozens of markers of interest individually in one experiment, so that rare EV subtypes (*e.g.* those containing tumor-derived mutated proteins) could be identified with reasonable certainty. Here we describe such a pipeline for antibody-based immuno sequencing (single EV immuno sequencing; seiSEQ) and which is able to result in readouts from single EV. We used droplet microfluidics to encapsulate individual antibody-DNA labeled EV into droplets that contain barcoded beads. Optimizing multiple extension and amplification steps, we show that multiplexed single EV protein profiling is feasible.

RESULTS/DISCUSSION

A droplet microfluidic platform for seiSEQ

Isolated EV were first labeled with Ab-DNA and remaining unbound Ab-DNA was removed by size exclusion chromatography (Izon).[23] (Fig. 1A) Ab-DNA labeled EV were then encapsulated into droplets along with barcoded beads. After droplet encapsulation, multiple extension and amplification steps were sequentially performed to synthesize amplicons and which are then sequenced to determine the protein make-up of specific vesicles. The approach used different barcodes to define protein types (Ab-DNA_{BC}) and individual vesicle (Bead-DNA_{BC}).

One of the challenges with using Bead-DNA_{BC} is the often inefficient reaction within droplets as the DNA is immobilized on beads. We therefore used a technique to dissolve polyacrylamide cross linked beads by breaking disulfide bridges with dithiothreitol (DTT).[24] Once cleaved, these beads rapidly release barcode primers (<3 mins at 1mM DTT), increasing the reaction efficiency in droplets and achieving high loading (>90%) of a single bead per droplet. The Bead-DNA_{BC} consisted of three sequence regions: a complementary sequence to the Ab-DNA_{BC}, a unique molecular identifier (UMI), and three combinatorial shorter barcode regions {Bead-DNA_{BC} : Bead-(bc1'-bc2'-bc3')-UMI-a} made by a 3-step split-pool approach so that individual vesicles can be identified with high diversity through sequencing of amplicons (Fig. 1B).

Target-specific antibodies of interest were conjugated to Ab-DNA_{BC} sequences that were computationally generated to prevent any sequence overlap. We used the bioorthogonal *trans*-cyclooctene/tetrazine (TCO/Tz) click chemistry to rapidly and efficiently conjugate Ab-DNA at high yields.[23] The Ab-DNA_{BC} consisted of three generic sequence regions: a complementary sequence to bind to Bead-DNA_{BC} in the droplet (a/a*), the actual antibody defining barcode (bc), and a T7 promoter sequence. The role of the T7 promoter sequence was to enable a more efficient in vitro transcription (IVT) to amplify RNA while minimizing crosstalk from incompletely extended DNA products.

We used a four-channel microfluidic device to encapsulate single EV and beads into droplets. (Fig. 1A, Fig. S2) Barcoded beads, labeled EV, oil, and master mix for an extension step were introduced through different channels to form droplets. Beads were closely packed by designing a channel that is narrower (40 μm in width) than the size of the beads (60 μm in diameter) as this is known to achieve efficient single bead loading per droplet (>90%).[25] Using this droplet maker, we created 180 μm droplets that contained beads and EV in a master mix solution. Different EV encapsulation conditions were explored and then validated by taking into account the Poisson distribution as a function of flow rates, EV input concentration, and droplet volume. We aimed to achieve 0.1 EV per droplet and at this ratio, the Poisson distribution predicts that ~9% of all droplets will have a single EV.[23] As the single bead loading efficiency is more than 90%, we calculated that ~8.1% of droplets contain both a single EV and a single bead.

seiSEQ pipeline

The seiSEQ pipeline includes five steps: extension of the Ab-DNA_{BC} and bead-DNA_{BC}, in vitro transcription (IVT) of the extended product, purification of IVT generated RNA, reverse transcription (RT) of RNA to cDNA, and polymerase chain reaction (PCR) of the cDNA. (Fig. 1C) The first step is an extension of the Ab-DNA_{BC} and bead-DNA_{BC} in droplets. During this step, bead-DNA_{BC} are dissociated from the beads to efficiently hybridize to Ab-DNA_{BC}. This is essential to generate a single strand that contains all the necessary information. We incorporated IVT in the pipeline to achieve two goals, i) signal amplification for single EV readout and ii) removal of a potential source of crosstalk. Multiple RNA copies can be efficiently synthesized from the Ab-DNA_{BC} due to the incorporated T7 promoter sequence. After IVT, the original DNA template strands and incompletely extended DNA products were removed using DNase to minimize crosstalk. Once DNA was removed, the amplified RNA was purified using AMPure XP magnetic beads and then converted to cDNA using RT. Converted cDNA were amplified using PCR for sequencing library preparation.

Validation of amplicon synthesized for single EV profiling

To validate a given amplicon synthesis, we first performed qPCR with converted cDNA. (Fig. 2A) In one experiment, a total of 350 EV were individually encapsulated into droplets. Two positive control samples were processed with a different numbers of bulk EV; a negative control sample contained all reagents but no EV. Both single and bulk EV samples showed comparable amplifications. From this result, we determined that the number of cycles (C_t = 32) required to selectively amplify the product while minimizing primer dimer formation. This approach was then used for further experiments. The length of the amplicon (152bp) was identical for both bulk and droplet single EV. (Fig. 2B) The single EV amplicon was further investigated using Sanger sequencing. (Fig. 2C) The amplicon sequence matched to the template sequence design, confirming successful amplicon synthesis for single EV protein profiling. After confirmation, following experiments were done using next generation sequencing to identify individual amplicons.

Accuracy and specificity of single EV profiling

To evaluate the accuracy of the single EV profiling technology, crosstalk of reads was measured using next generation sequencing. (Fig. 3A) For this experiment, an anti-EGFR antibody was conjugated to two different DNA barcode sequences. Both synthetic Ab-DNA_{BC} were used to separately label Gli36-glioma cell line derived EV. Labeled EV were then mixed prior to droplet encapsulation. The developed pipeline was used to synthesize sequencing amplicons and the sequencing data was aligned to each barcode sequence to measure crosstalk reads. A majority of the reads was correctly aligned to one barcode sequence or the other as one would expect. There was no crosstalk and we only observed 5% cross contamination.

To evaluate the specificity of seiSEQ, another control experiment was performed to compare the number of reads obtained from isotype control antibody labeled EV to that from target specific antibody labeled EV. (Fig. S3) Due to the scarcity of the protein molecules from individual EV, it is important to find a threshold that can distinguish a target specific signal

from a nonspecific binding signal. To set a threshold, gli36-glioma cell line derived EV were labeled with both anti-IgG isotype control antibody and anti-EGFR antibody and single EV were sequenced. Histograms were created and threshold was drawn at a 95% confidence interval of the reads from anti-IgG isotype antibody-DNA. For example, 95% of the EV that were labeled with anti-IgG isotype antibody-DNA resulted in 0, 1, or 2 reads and only the EV that have more than 2 reads were analyzed from the anti-EGFR antibody-DNA labeled sample. The same approach was used for future analysis.

seiSEQ of macrophage derived vesicles

To show the potential for multiplexed single EV analysis, we performed next generation sequencing on 8 proteins (CD9, F4/80, CD11b, CD63, CD45, CD81, two isotype controls) in 1100 EV obtained from RAW 264.7 murine macrophage cell line. (Fig. 3B) The results show that the majority of EV had a low number of proteins of interest with CD9 being the most abundant one (89% of EV had this protein) followed by CD81+ (25%), CD63+ (12%), CD11b+ (8.9%), F4/80+ (2.1%), and CD45+ (1.1%) populations. These results are not entirely unexpected as CD9 is a canonical marker of EV.[26] We next determined the EV protein co-expression levels. (Fig. 3C) This is important because co-positive populations could provide molecular information on EV subtypes that cannot be derived by bulk measurements. The results show co-positive populations of RAW264.7 EV that were comprised of CD9+CD81+ (17.9%), CD9+CD63+ (9.4%), CD63+CD81+ (3.1%), CD9+CD63+CD81+ (2.7%), F4/80+CD11b+ (0.27%), F4/80+CD45+ (0.09%), and no F4/80+CD11b+CD45+ populations. A final experiment was performed to compare seiSEQ results from single EV to bulk measurements as this can be done mathematically by summing the seiSEQ data. (Fig. 3D) These results showed good correlation (R^2=0.85) between bulk EV measurements measured by flow cytometry and single EV measurements by seiSEQ. Cell expression levels measured by flow cytometry were included as a positive control.

CONCLUSIONS

Here we present an antibody-DNA barcode based immuno sequencing method (seiSEQ) that allows multiplexed measurement of proteins on nanomaterials. For proof of principle studies, we analyzed EV from mammalian cells as these are an increasingly recognized as potentially useful biomarkers (“liquid biopsy”). One of the clinical challenges is to detect cancers much earlier than is currently possible and single EV analytical techniques are expected to play an important role in this application. To provide specificity, we hypothesized that either the detection of rare mutated proteins (*e.g.* KrasG12D) or the co-expression pattern of ubiquitously expressed proteins (*e.g.* EGFR+/EPCAM+/HER2+) could ultimately provide a way to determine whether a vesicle was shed by a tumor cell. To achieve the goal of multiplexing beyond what is possible by fluorescence imaging (often < 2–3 channels given the small size and needs for orthogonally compatible amplification strategies), we used droplet microfluidics to compartmentalize and barcode individual EV and sequencing to derive the protein composition of individual vesicles. The use of droplets allows one to continuously generate compartmentalized reaction chambers to rapidly perform chemical reactions in confined droplet spaces. This provides an improvement in

throughput over an analogous approach in microwell plates, which also requires antibody-immobilization.[27]

A number of multiplexing technologies have been described, differing in sensitivity, scale, and throughput. For single EV analyses, most optical based methods have limited multiplexing capabilities because of spectral overlap and ambiguities in such small scale. The seiSEQ method described here allows DNA-based amplification, which is robust and highly sensitive. Second, the method has a nearly unlimited multiplexing capability based only by DNA barcode design. Finally, the seiSEQ method can be used to profile large numbers of EV (ultra-high throughput) when combined with deep sequencing. These three attributes make the seiSEQ method advantageous and scalable.

To validate the sequencing protocol and show its ability to profile single EV, we initially chose a cost effective sequencing service ($75/sample, Amplicon-EZ, Genewiz) instead of full sequencing. The service provides ~50,000 reads per sample, which allowed us to profile ~1,100 EV. The ultimate throughput is not limited by the seiSEQ technology per se but rather by the number of reads from sequencing. The seiSEQ technology can be used in high throughput mode when combined with conventional sequencers (*e.g.* HiSeq, NextSeq, *etc*). The latter allows one to profile more than 10^5–10^6 EV, although at higher costs.

Depending on the type of EV, their volume is approximately 10^6 fold smaller than that of a mammalian cell, limiting the number of even “abundant” proteins. This and the stochastic processes of protein distribution result in vesicular protein concentrations at the (sub-)femptomolar level. The current platform could be further improved to allow the detection of even more scarce proteins. For example, to increase the signal-to-noise (SNR) ratio, methods such as the proximity ligation assay (PLA) could be used to minimize background derived from remaining unlabeled Ab-DNA or different blocking buffers could be explored in an effort to minimize nonspecific binding. Additionally, the measurement accuracy could be increased by increasing sequencing depth. This may be possible by sequencing the same region multiple times and with a higher number of reads. Due to the scarcity and limited number of EV proteins, millions of reads may be sufficient to sequence the sample in depth. These advances will potentially enable a further improved measurement accuracy of seiSEQ.

While we focused on EV, the seiSEQ technique could also be used to profile rare proteins on other nanoparticles. This for example provides the possibility of profiling viral coat proteins, bacterial proteins, and proteins coated on synthetic nanomaterials (corona). We anticipate that the seiSEQ will become a versatile tool to profile rare and diverse subpopulations of bionanomaterials. This would be useful in biomarker discovery and the development of medical diagnostics and therapeutics.

METHODS/EXPERIMENTAL

Device fabrication

The microfluidic device for droplet generation was fabricated at the Soft Materials Cleanroom (SMCR), Harvard Center for Nanoscale Systems (CNS). The device (*h*

= 100 μm) was made using soft lithography with SU-8 3050. The PDMS that consists of microfluidic channels were bonded with glass using plasma bonding. The device was made hydrophobic before usage by treating with 1% Trichloro(1H,1H,2H,2H-perfluorooctyl)silane in Novec 7500 (Oakwood Chemical).

Cell culture and EV isolation

Gli36wt, gli36vIII, and RAW264.7 cell lines were used to test and optimize the seiSEQ technology. Cells were grown in a 150mm cell culture dish and expanded to 8–12 dishes for EV collection. Cells were grown and passaged in DMEM (10% FBS, 1% Penicillin/Streptomycin). Once confluent, media was changed to exosome-depleted DMEM (5% exosome-depleted FBS, 1% Penicillin/Streptomycin) and supernatant was collected 48 hours after the media change. The collected supernatant was spun at 400g for 5 mins and filtered with 0.22 μm vacuum filterto remove any cellular debris. Then, the supernatant was centrifuged (Beckman Coulter) at 100,000 for 70 mins at 4°C for two times. The EV pellet was resuspended in PBS and aliquoted and stored in −80C until usage.

Antibodies

Cetuximab (anti-EGFR antibody, Erbitux), anti-CD63 antibody (Ancell, 215–820), anti-CD9 antibody (BioLegend, cat#124802), anti-CD63 antibody (BioLegend, cat#143901), anti-CD81 antibody (BioLegend, cat#104901), anti-F4/80 antibody (BioXCell, BE0206), anti-CD11b antibody (BioXCell, BE0007), anti-CD45 antibody (R&D Biosystems, MAB114), rat IgG2a isotype control (BioXCell, BE0089), and rat IgG2b isotype control (BioXCell, BE0090) were used to test and optimize the technology. All antibodies were tested on positive cell lines and validated before usage, checked for the absence of BSA for Ab-DNA conjugation, and conjugated with AFDye 647 NHS Ester (Click Chemistry Tools, cat#1344) for flow cytometry.

EV characterization (Qubit, NTA)

After isolation, EV was characterized in two different ways. The protein concentration was measured using Qubit (Thermo Fisher) and the number of particles was calculated using nanoparticle tracking analysis (NTA). For Qubit, the protein assay kit (Thermo Fisher) was used and the company protocol was followed for measurement. For NTA, the measurement was done at the Nanosight Nanoparticle Sizing and Quantification Facility at Massachusetts General Hospital (MGH). Three 30 sec measurements were performed and averaged from each sample. The same parameters were used for analysis (Image: Screen gain of 7.4, Camera level of 11, Detection: Screen gain of 10, Detection threshold of 13)

EV labeling and purification

EV was labeled with 10μg/ml of Ab-DNA conjugates in 1% BSA-PBS for 1 hour with mixing and purified using size exclusion chromatography, qEV column (Izon science), to remove unlabeled Ab-DNA conjugates. Single use qEV column was used and 400μl was collected after dead volume to achieve a pure EV population. The labeled EV was stored in 4°C until usage and used within a few days to prevent degradation.

Barcoded bead fabrication

500 μL solution mix was prepared containing 50 μL TBSET buffer, 30 μL 10% (w/v) APS (Sigma-Aldrich, A9164), 75 μL 40% (v/v) Acrylamide solution (Sigma-Aldrich, A4058–100ML), 20 μL 250 μM Acrydite-modified DNA primers (IDT, sequence in Table S1), 245 μL 0.8% (w/v) BAC (Sigma-Aldrich, A4929–5G) and 80 μL H_2O. This solution was loaded into a 1 mL syringe (Becton Dickinson, 309628). 1.5 mL carrier oil (RAN Biotechnologies, 008-FluoroSurfactant-2wtH-50G) and 6 μL of TEMED (Sigma-Aldrich, T9281–25ML) were mixed and loaded into a 3-mL syringe (Becton Dickinson, 309657). These two syringes were connected with inlets of the droplet generation device (Fig S1) by PE2 tubing (Scientific Commodities, BB31695-PE/2). The aqueous solution was run at 500 μL/hr and the oil at 1000 μL/hr. The emulsion droplets were then collected from the outlet of the microfluidics chip. The collected droplets were covered with 200 μL mineral oil (Sigma-Aldrich, M5310–1L) and incubated at 70°C overnight. The carrier oil and mineral oil phases were centrifuged and discarded. 500 μL 20% (vol/vol) PFO (Alfa Aesar, B20156) in HFE 7500 (Novec 7500) was used to break the droplets. The beads in the aqueous phase were washed with 1% Span-80 (Sigma-Aldrich, S6760–250ML) in hexane (Sigma-Aldrich, 227064–1L) twice and then with TBSET buffer 3 times. The beads were filtered using a 70 μm cell strainer (Corning, 352350) and then stored in TET buffer at 4°C for up to 6 months.

DNA barcodes

Two types of DNA barcodes were used in this study. First, DNA barcodes for beads were synthesized using a 3-step extension. Acrydite DNA was used to make acrylamide-based hydrogel beads and barcodes were extended for three times with 96 primer diversity each time to achieve high throughput EV profiling. The DNA barcodes for antibodies consist of three regions - T7 promoter sequences for IVT, barcode sequence, and a universal sequence complementary to the sequence of barcoded beads.

seiSEQ protocol

EVs were first isolated from plasma or cell cultured media using ultracentrifugation or size exclusion chromatography. Isolated EVs were labeled with antibody-DNA conjugates and purified using size-exclusion chromatography to remove unbound antibody-DNA conjugates. Labeled EVs were then encapsulated into droplets (0.1 EV per droplet using Poisson distribution) along with barcoded beads and master mix (19.2μl 10mM dNTP, 6.48μl 10% triton, 14.4μl 100mM DTT, 14.4μl 10x TP, 5.76μl BST 2.0 warmstart, and 4.32μl USER enzyme). With the collected droplets, we performed an extension step (60°C for 2hr) using a thermal cycler. We then broke the droplets using PFO and collected the upper phase for IVT using the MEGAshortscript T7 Transcription kit (Thermo Fisher). We purified RNA using an AMPure bead (Beckman Coulter) with 1.6x volume of the sample and eluted the sample in RNA elution buffer. We performed reverse transcription (RT) using Maxima H Minuse Reverse Transcriptase (Thermo Fisher). After RT, we set up for PCR and prepare for sequencing. Sequencing was performed using a Next Gen Sequencing service from the Genewiz company.

Ab-DNA conjugation

BSA free antibodies were buffer exchanged to biocarbonate buffer (pH8.4) using a 40k Zeba column (Thermo Fisher, 87765). The antibody was incubated with TCO-PEG4-NHS Ester (Click Chesmitry Tools, A137–10) for 25 mins at room temperature and unlabeled TCO-PEG4-NHS Ester was removed using a 40k Zeba column. Degree of labeling (DOL) was checked by incubating antibodies with Cy3 Tetrazine (Click Chemistry Tools, 1018–1) for 25 mins at room temperature and remaining Cy3 Tetrazine was removed using a 40k Zeba column. Cy3:Antibody ratio was measured using the Nanodrop UV/Vis mode (Thermo Scientific) at A550/A280.

1mM of amine-modified DNA oligo (IDT) was exchanged to borate buffer (pH8.5) using a 7k Zeba column (Thermo Fisher, 89878). The DNA oligo was incubated with Methyltetrazine-PEG4-NHS Ester (Click Chemistry Tools, 1069–10) for 25 mins at room temperature and unlabeled Tz-PEG4-NHS was removed using three 7k Zeba columns. Tz:DNA ratio was measured using the Nanodrop UV/Vis mode at A520/A260. TCO labeled antibody and Tz labeled DNA were mixed with appropriated DNA excess (Cy3:Antibody ratio - 0.5) and incubated for 45 mins at room temperature. The conjugation was validated using the NuPAGE 4–12% Bis-Tris Protein Gel (Thermo Fisher, NP0321BOX). Unconjugated antibody and DNA-conjugated antibody were incubated with 4x NuPAGE LDS Sample Buffer (Thermo Fisher, NP0007) for 5 mins at 75°C and loaded to the gel with Novex Sharp Pre-stained Protein Standard (Thermo Fisher, LC5800). The gel was run in 20x NuPAGE MOPS SDS Running Buffer (Thermo Fisher, NP0001) for 1 hour at 120V. The validated antibody-DNA conjugate was stored in 4°C until usage.

Flow cytometry

Cells were incubated with 5 μg/ml of antibodies in 1% BSA-PBS at 4°C for 20 mins and washed twice. EV were mixed with 4 μm aldehyde/sulfate latex beads (Thermo Fisher, A37304) in PBS and incubated for 2 h at room temperature or overnight at 4°C with rotation. Glycine was added at 100mM final concentration and incubated for 30 mins at room temperature. EV captured beads were centrifuged for 3 mins at 4000rpm. The pellet was resuspended in 0.5% BSA-PBS and washed twice. Beads were incubated with 5 μg/ml of antibodies in 0.5% BSA-PBS for 30 mins at 4°C with rotation. Beads were then washed twice with 0.5% BSA-PBS. The LSR II Flow Cytometer (BD Biosciences) was used for measurements and the FlowJo program was used for data analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported in part by the following NIH grants: PO1 CA069246, RO1 CA204019, R21 CA236561. J.K. is supported by the Schmidt Science Fellow program, in partnership with the Rhodes Trust, Oxford, UK. Y.W. is supported by a QuantBio graduate student award at Harvard.

REFERENCES

1. Nilsson J; Skog J; Nordstrand A; Baranov V; Mincheva-Nilsson L; Breakefield XO & Widmark A. Prostate Cancer-Derived Urine Exosomes: A Novel Approach to Biomarkers for Prostate Cancer. Br J Cancer. 2009, 100, 1603–1607. [PubMed: 19401683]
2. Graner MW; Alzate O; Dechkovskaia AM; Keene JD; Sampson JH; Mitchell DA & Bigner DD Proteomic and Immunologic Analyses of Brain Tumor Exosomes. FASEB J. 2009, 23, 1541–1557. [PubMed: 19109410]
3. Mathivanan S; Lim JW; Tauro BJ; Ji H; Moritz RL & Simpson RJ Proteomics Analysis of A33 Immunoaffinity-Purified Exosomes Released from the Human Colon Tumor Cell Line Lim1215 Reveals a Tissue-Specific Protein Signature. Mol Cell Proteomics. 2010, 9, 197–208. [PubMed: 19837982]
4. Théry C. Cancer: Diagnosis by Extracellular Vesicles. Nature. 2015, 523, 161–162. [PubMed: 26106856]
5. Colombo M; Raposo G. & Théry C. Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. Annu Rev Cell Dev Biol. 2014, 30, 255–289. [PubMed: 25288114]
6. Schorey JS & Harding CV Extracellular Vesicles and Infectious Diseases: New Complexity to an Old Story. J Clin Invest. 2016, 126, 1181–1189. [PubMed: 27035809]
7. EL Andaloussi S; Mäger I; Breakefield XO & Wood MJ Extracellular Vesicles: Biology and Emerging Therapeutic Opportunities. Nat Rev Drug Discov. 2013, 12, 347–357. [PubMed: 23584393]
8. Raposo G. & Stoorvogel W. Extracellular Vesicles: Exosomes, Microvesicles, and Friends. J Cell Biol. 2013, 200, 373–383. [PubMed: 23420871]
9. Eitan E; Green J; Bodogai M; Mode NA; Bæk R; Jørgensen MM; Freeman DW; Witwer KW; Zonderman AB; Biragyn A; Mattson MP; Noren Hooten N. & Evans MK Age-Related Changes in Plasma Extracellular Vesicle Characteristics and Internalization by Leukocytes. Sci Rep. 2017, 7, 1342. [PubMed: 28465537]
10. Lee K; Fraser K; Ghaddar B; Yang K; Kim E; Balaj L; Chiocca A; Breakefield XO; Lee H. & Weissleder R. Multiplexed Profiling of Single Extracellular Vesicles. ACS Nano. 2018, 12, 494–503. [PubMed: 29286635]
11. Fraser K; Jo A; Giedt J; Vinegoni C; Yang KS; Peruzzi P; Chiocca EA; Breakefield XO; Lee H. & Weissleder R. Characterization of Single Microvesicles in Plasma from Glioblastoma Patients. Neuro Oncol. 2019, 21, 606–615. [PubMed: 30561734]
12. Shao H; Im H; Castro CM; Breakefield XO; Weissleder R. & Lee H. New Technologies for Analysis of Extracellular Vesicles. Chem Rev. 2018, 118, 4, 1917–1950. [PubMed: 28443658]
13. Im H; Shao H; Weissleder R; Castro CM & Lee H. Nano-Plasmonic Exosome Diagnostics. Expert Rev Mol Diagn. 2015, 15, 725–733. [PubMed: 25936957]
14. Im H; Lee K; Weissleder R; Lee H. & Castro CM Novel Nanosensing Technologies for Exosome Detection and Profiling. Lab Chip. 2017, 17, 2892–2898. [PubMed: 28745363]
15. Prada I; Amin L; Furlan R; Legname G; Verderio C. & Cojoc D. A New Approach to Follow a Single Extracellular Vesicle-Cell Interaction Using Optical Tweezers. Biotechniques. 2016, 60, 35–41. [PubMed: 26757810]
16. Gualerzi A; Niada S; Gannasi C; Picciolini S; Morasso C; Vanna R; Rossella V; Masserini M; Bedoni M; Ciceri F; Ester Bernardo M; Teresa Brini A. & Gramatica F. Raman Spectroscopy Uncovers Biochemical Tissue-Related Features of Extracellular Vesicles from Mesenchymal Stromal Cells. Sci Rep. 2017, 7, 9820. [PubMed: 28852131]
17. Smith ZJ; Lee C; Rojalin T; Carney RP; Hazari S; Knudson A; Lam K; Saari H; Lazaro Ibañez E; Viitala T; Laaksonen T; Yliperttula M. & Wachsmann-Hogiu S. Single Exosome Study Reveals Subpopulations Distributed among Cell Lines with Variability Related to Membrane Content. J Extracell Vesicles Journal of Extracellular Vesicles. 2015, 4, 10.3402/jev.v4.28533.
18. Löf L; Ebai T; Dubois L; Wik L; Ronquist KG; Nolander O; Lundin E; S√∂derberg O; Landegren U. & Kamali-Moghaddam M. Detecting Individual Extracellular Vesicles Using a Multicolor in

Situ Proximity Ligation Assay with Flow Cytometric Readout. Scientific Reports. 2016, 6, 34358. [PubMed: 27681459]

19. Klein AM; Mazutis L; Akartuna I; Tallapragada N; Veres A; Li V; Peshkin L; Weitz DA & Kirschner MW Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell. 2015, 161, 1187–1201. [PubMed: 26000487]
20. Macosko EZ; Basu A; Satija R; Nemesh J; Shekhar K; Goldman M; Tirosh I; Bialas AR; Kamitaki N. & Martersteck EM Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015, 161, 1202–1214. [PubMed: 26000488]
21. Weissleder R. & Pittet MJ The Expanding Landscape of Inflammatory Cells Affecting Cancer Therapy. Nat Biomed Eng. 2020, 4, 489–498. [PubMed: 32203281]
22. Gyuris A. Navarrete-Perea J; Jo A; Cristea S; Zhou S; Fraser K; Wei Z; Krichevsky AM; Weissleder R; Lee H; Gygi SP & Charest A. Physical and Molecular Landscapes of Mouse Glioma Extracellular Vesicles Define Heterogeneity. Cell Rep. 2019, 27, 3972–3987.e6. [PubMed: 31242427]
23. Ko J; Wang Y; Carlson JCT; Marquard A; Gungabeesoon J; Charest A; Weitz D; Pittet MJ & Weissleder R. Single Extracellular Vesicle Protein Analysis Using Immuno-Droplet Digital Polymerase Chain Reaction Amplification. Advanced biosystems. 2020, 4, 1900307.
24. Wang Y; Cao T; Ko J; Shen Y; Zong W; Sheng K; Cao W; Sun S; Cai L. & Zhou Y. Dissolvable Polyacrylamide Beads for High-Throughput Droplet DNA Barcoding. Advanced Science. 2020, 7, 1903463.
25. Abate AR; Chen C-H; Agresti JJ & Weitz DA Beating Poisson Encapsulation Statistics Using Close-Packed Ordering. Lab on a Chip. 2009, 9, 2628–2631. [PubMed: 19704976]
26. Yoshioka Y; Konishi Y; Kosaka N; Katsuda T; Kato T. & Ochiya T. Comparative Marker Analysis of Extracellular Vesicles in Different Human Cancer Types. Journal of extracellular vesicles. 2013, 2, 20424.
27. Wu D; Yan J; Shen X; Sun Y; Thulin M; Cai Y; Wik L; Shen Q; Oelrich J. & Qian X. Profiling Surface Proteins on Individual Exosomes Using a Proximity Barcoding Assay. Nature communications. 2019, 10, 1–10.

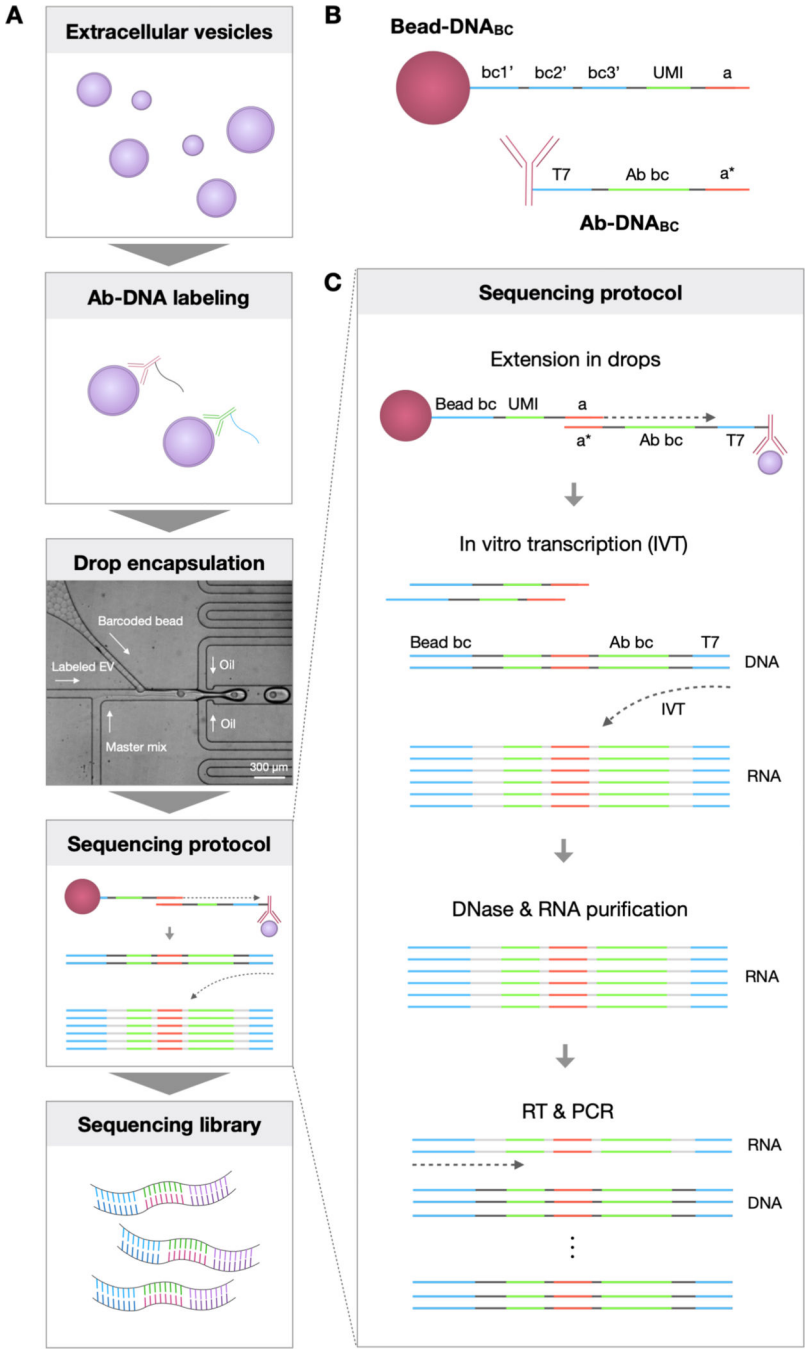

Fig 1. Schematic of seiSEQ.

A) The pipeline includes EV labeling with Ab-DNABC constructs, drop encapsulation with barcoded beads, and single EV sequencing protocol. The drop encapsulation step includes a microscopic image of a droplet generator with four input channels for oil, barcoded beads, labeled EV, and master mix, and one output channel to collect individual droplets (scale bar = 300 μm). B) DNA sequence composition on barcoded beads (bc1', bc2', bc3' = three subbarcoded regions created using a split-pool approach during bead synthesis; UMI = unique molecular identifier; a= hybridizing sequence to Ab-DNABC) and

antibodies (T7 = T7 promoter sequence; Ab bc = antibody barcode; a* = complementary strand to ‘a’ on the bead-DNABC). See Fig S2 for details. C) Schematic on the sequencing protocol. Ab-DNABC and Bead-DNABC are hybridized at the a/a* sequence region. After hybridization, extension is performed within droplets. The extended product consists of bead barcode (Bead bc), antibody barcode (Ab bc), UMI, and T7 promoter sequence. The T7 promoter sequence is used to efficiently amplify RNA. Then, DNase is treated to remove any remaining DNA and RNA is purified and converted to cDNA using RT. The cDNA undergoes PCR for amplification and post-PCR purification for sequencing.

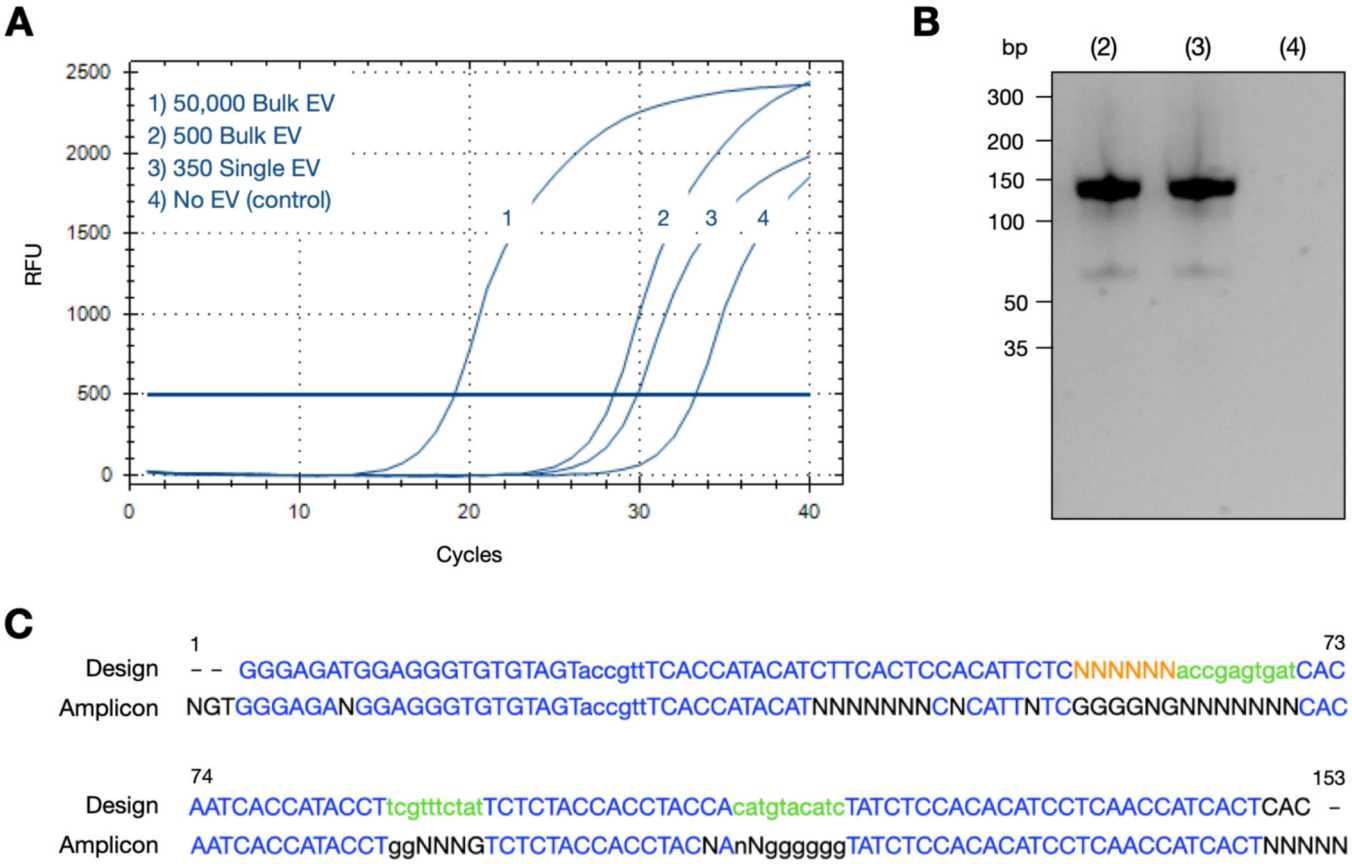

Fig 2. Amplicon generation and validation.

A) Shown are the qPCR cycle curves for 4 different samples: (1) bulk 50,000 EV, (2) bulk 500 EV, (3) single EV amplicons from 350 EV and generated using the protocol shown in Fig 1; and (4) negative control containing all the reagents but no EV. B) The PCR amplified samples (2, 3, 4) with Ct = 32 were run on a gel to determine amplicon size. Two main bands were observed at ~150 bp for (2) and (3), a size that matches to the template sequence (152 bp). As expected, there was no band in the control sample (4) in which EV were missing. C) Sanger sequencing was performed to compare the amplicon sequence from sample (3) to the original template sequence design. The sequence of the amplicon matched the template sequence design (blue: matching sequences, orange: UMI, green: bead/EV barcodes, black: unmatched sequences that are mostly in barcode regions). Using Sanger sequencing, barcode regions were expected not to be matched to a specific sequence as they vary from one amplicon to the other. Following experiments were done with next generation sequencing to identify individual amplicons.

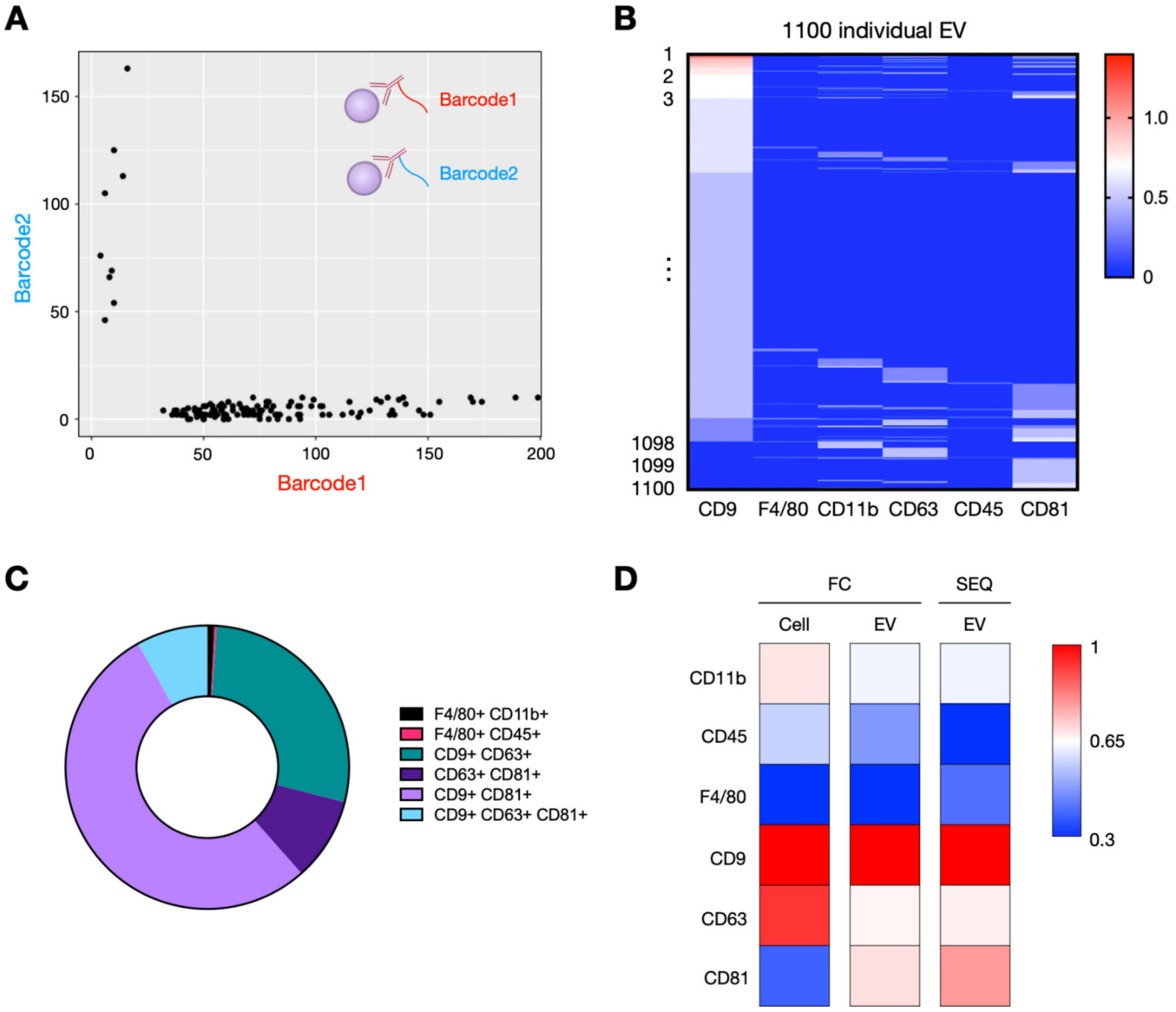

Fig 3. seiSEQ of macrophage derived vesicles.

A) Crosstalk analysis. To determine potential crosstalk between different EV, two barcodes (barcode 1, 2) were used to separately label EV and which were then mixed for analysis. In the example shown there was no crosstalk. B) Heatmap of 6 different protein markers in 1,100 individual EV. CD9 was the most abundant being present in 89% of vesicles, followed by CD81+ (25%), CD63+ (12%), CD11b+ (8.9%), F4/80+ (2.1%), and CD45+ (1.1%). C) A ring plot of co-positive marker populations in macrophage EV. They were comprised of F4/80+CD11b+ (0.27%), F4/80+CD45+ (0.09%), CD9+CD63+ (9.4%), CD63+CD81+ (3.1%), CD9+CD81+ (17.9%), CD9+CD63+CD81+ (2.7%), and no F4/80+CD11b+CD45+ populations. D) Comparison of bulk EV measurements using flow cytometry (FC) and seiSEQ. FC was used to profile the shown biomarkers in bulk samples. seiSEQ was used for single EV (shown are the summary results from 1,100 single EV). Note the good correlation.