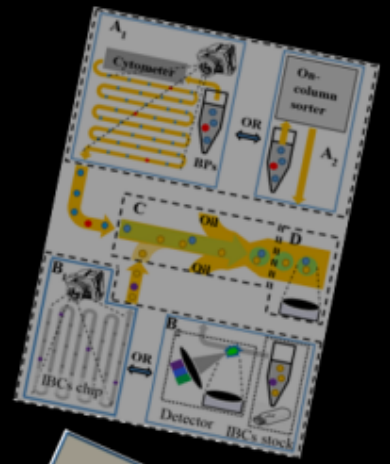
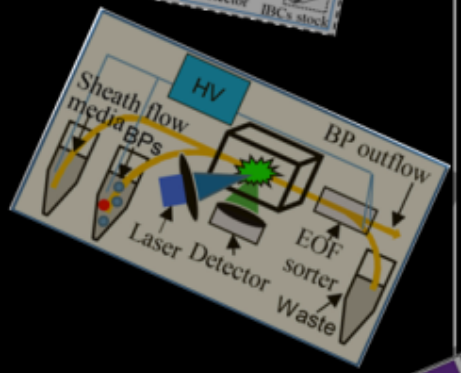
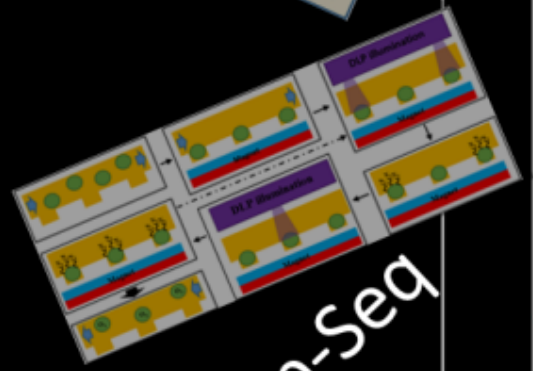
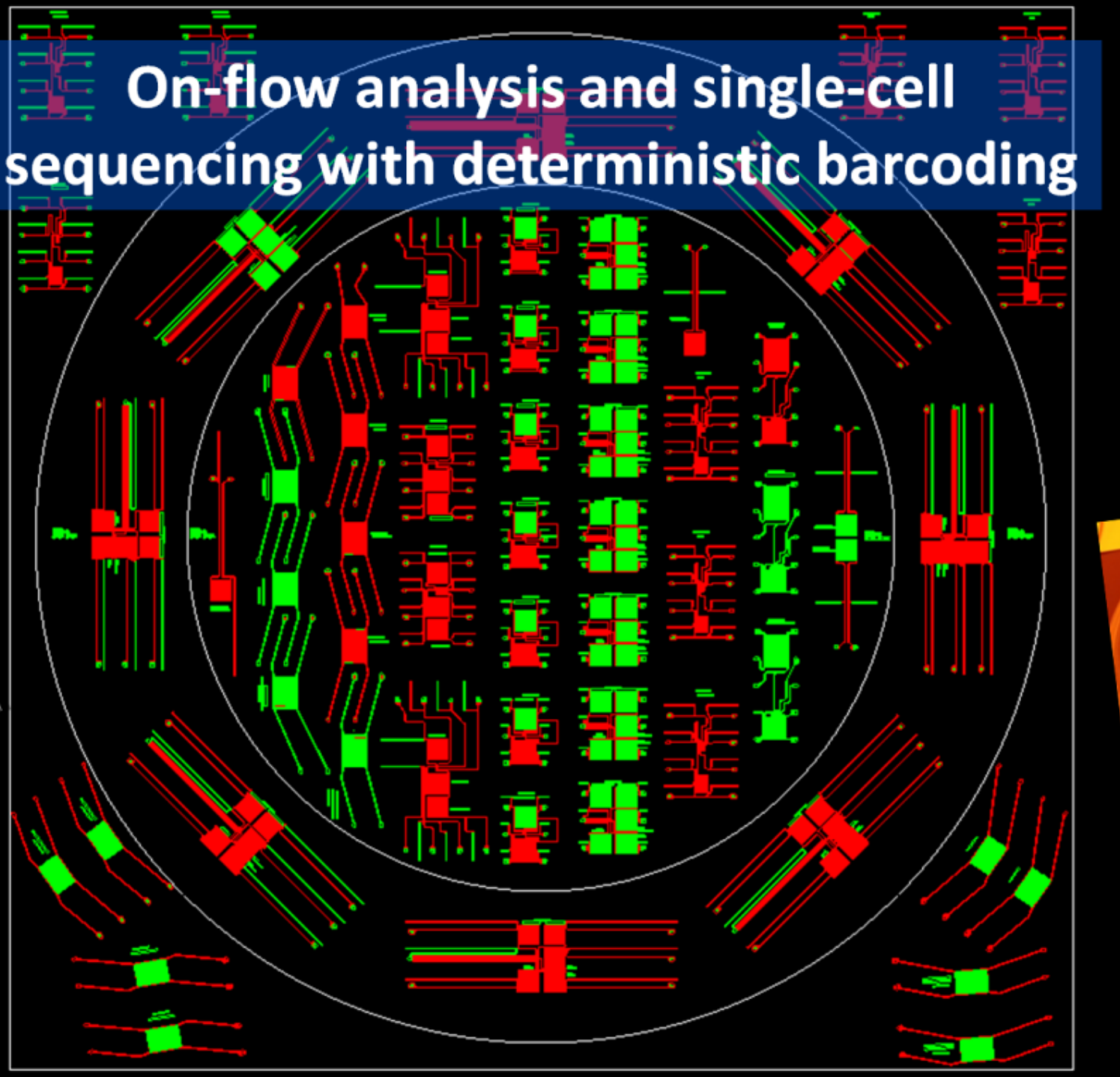
On-flow analysis and single-cell sequencing with deterministic barcoding

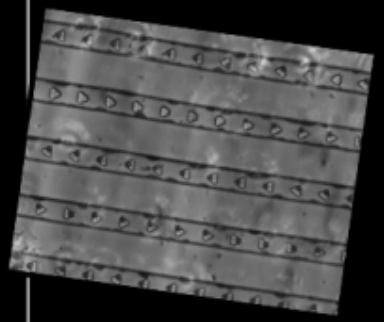


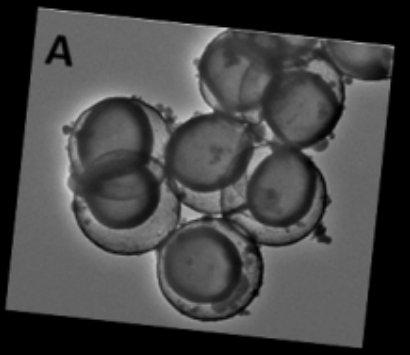




FACS + Drop-Seq







High-throughput analysis of individual particles vs analysis in-bulk.

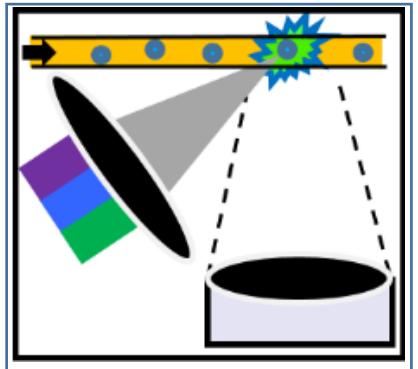
- Heterogeneity is a key for understanding development and aging, fertility, carcinogenesis, selective therapies, resistance and infectability

- Analytical paradigm shifts as soon as methods allow – “averaged temperature over a hospital” is of little use!

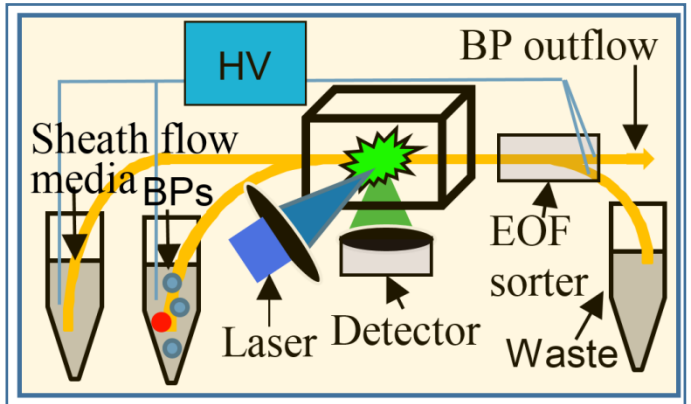
- Two little revolutions: Cytometry and Drop-Seq

On-flow analysis: flow cytometry, FACS and relatives

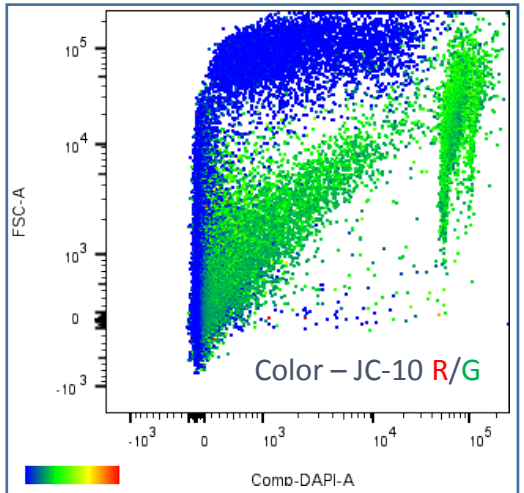
- **Particles:** cells (cytometry, FACS), organelles, exosomes, viruses, etc.
- **Detection:** on-column, sheath-flow
- **Multiple channels of analysis**: angle-dependent and multi-wave scattering, fluorescence, polarization, SERS, migration time, etc.
- **Driving:** pressure, electroosmotic flow, electrophoresis, ultrasound, laser pumping, etc.
- **Sorting:** in-droplets, in-flow
- **High-throughput, non-destructive**



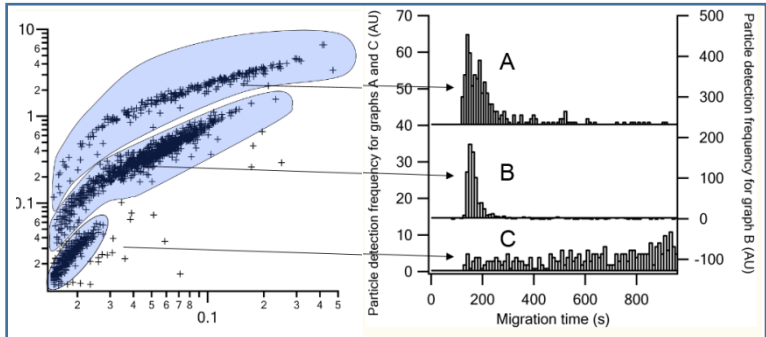
Example: on-flow detection



Example: EOF-driven sorter



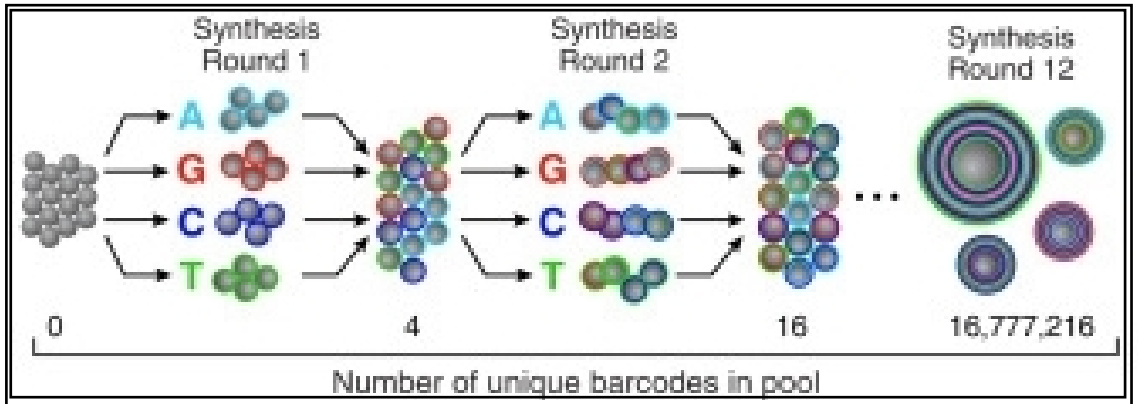
Example: four-channel analysis, DNA content vs. size vs. membrane potential



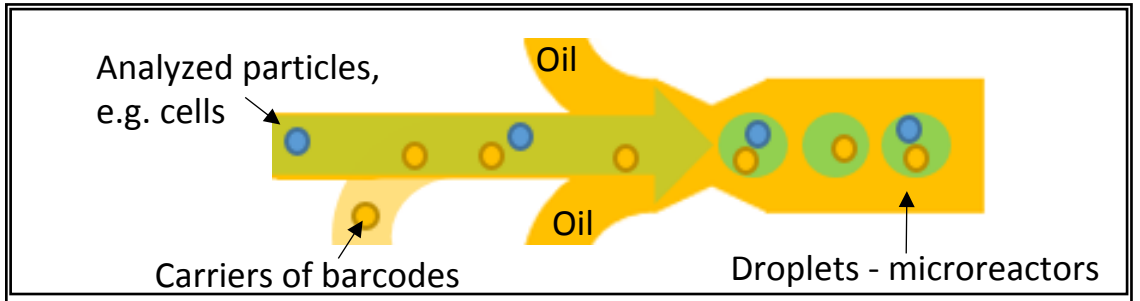
Example: fluorescence, scattering and electrophoretic mobility

High-throughput individual particle sequencing: Drop-Seq and relatives

- Make carriers of unique oligonucleotide barcodes.
- Pair analyzed particles with barcode carriers in individual microreactors.
- Ligate genetic content with barcodes
- Pool, sequence, identify sequences by barcodes



Split-and-pool synthesis of barcodes *



Pairing of analyzed biological particles with genetic content (**BPs**) with carriers of barcoding oligonucleotides.

* Lan, F., et al., *Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding.* Nat Biotechnol, 2017. 35(7): p. 640-646.

Phenotype + genome/transcriptome: the third revolution?

- **On-flow analysis, e.g. cytometry**: non-destructive, high-throughput methods, providing comprehensive data BPs* phenotypes and functionality.

- **Individual cell sequencing, e.g. Drop-Seq** – high-throughput methods providing comprehensive data on BPs genotype and transcriptome.

- **Combination** of these two would simultaneously analyze phenotype, functionality, genotype and transcriptome of each BP in a large community**.

* Biological Particle with genetic content, such as somatic and germ cells, pathogenic and symbiotic microbiota, mitochondria, nuclei, nucleoli, ribosomes, viruses, and exosomes.
** Biome, tissue, culture, totality of mitochondria in a cell, etc.

Phenotype + genome/transcriptome: the third revolution?

Examples of affected fields:

- cancer development, such as polyclonal seeding in metastatic cancers
- prevention of carcinogenesis in stem cells therapies
- therapy resistance development
- manifestation of genome in variable transcriptomes and phenotypes during development
- pathogen-cell interactions, infectability
- complex microbial communities and their interactions with host
- mitochondrial heteroplasmy manifestation in mitochondrial tagging for mitophagy, replication or initiation of apoptosis, in compromising of apoptotic antiproliferation, in neural and muscular degeneration, infertility, inherited pathologies and mitochondrial community rejuvenation between generations

Phenotype + genome/transcriptome: the third revolution?

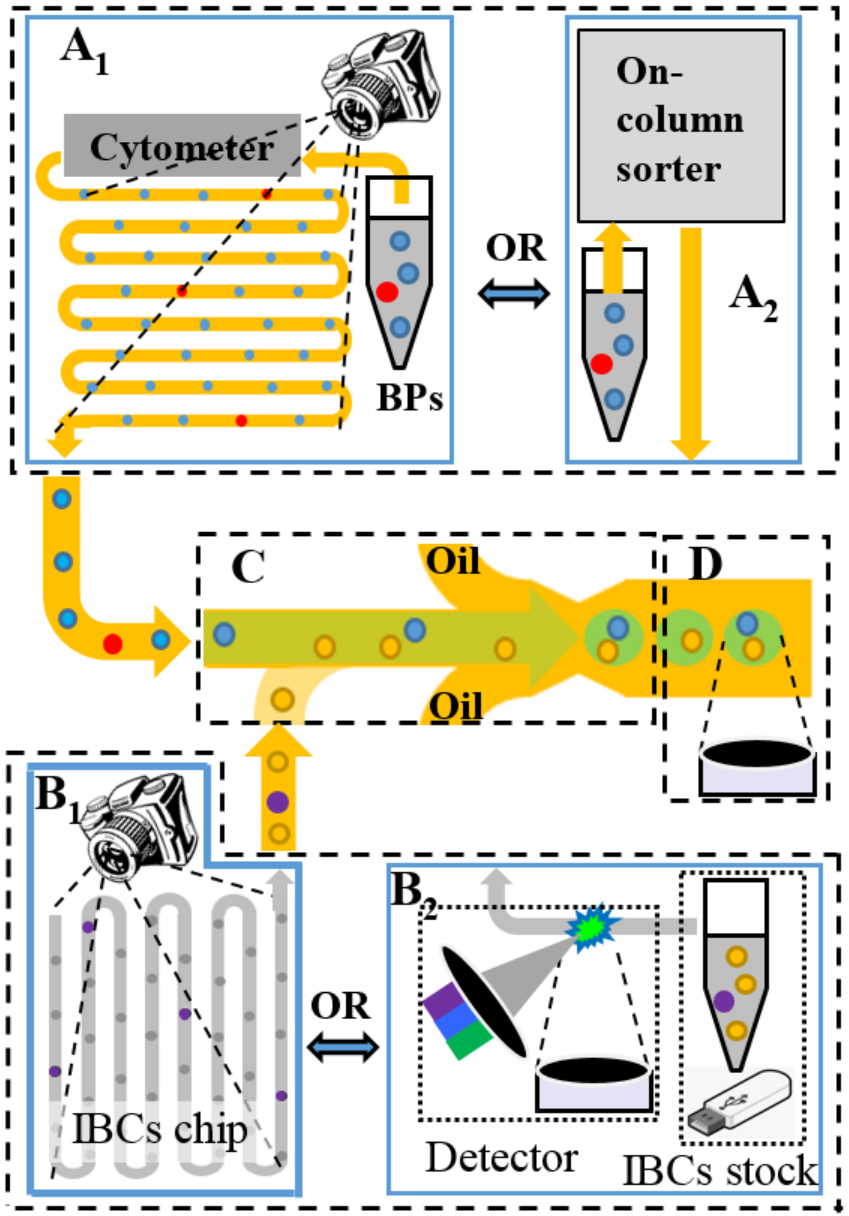
- **CITE-SEQ*: the first approach.** Antibody-oligonucleotide tagging enables limited phenotype analysis using cell surface markers. Excellent, but limited comparing to flow cytometry.

- **Deterministic barcoding** – our attempt of the integration. We employ identifiable particles carrying unique and known oligonucleotide barcodes to micro droplets of Drop-Seq, as opposed to the unique but random barcodes, used today.
- Identifiable barcoding carriers (IBC), serially paired with BPs in a cytometer outflow in a regular Drop-Seq chip would maintain the desired links between cytometric and Drop-Seq data

* Stoeckius, M., et al., *Simultaneous epitope and transcriptome measurement in single cells.* Nat Methods, 2017. **14**(9): p. 865-868.

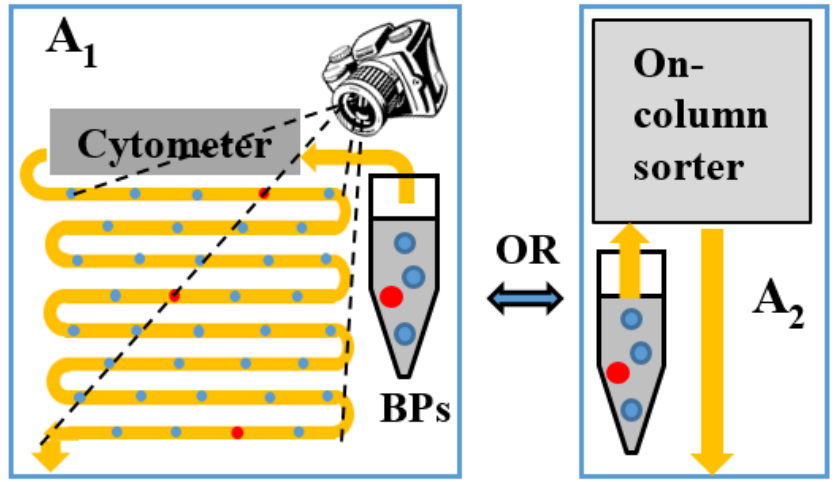
Cytometry + Drop-Seq via deterministic barcoding

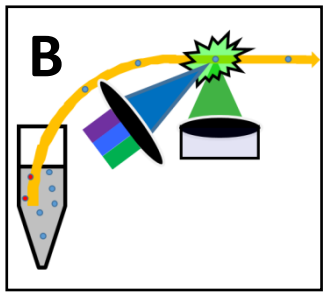
- On-flow analysis module (either A_1 or A_2) – a cytometer in the first approximation - supplies stream of analyzed BPs to the droplet generator (C).
- Simultaneously, IBCs (from either B_1 or B_2 source) are supplied to the same generator.
- Reference particles (red BPs, violet IBCs) in droplets are monitored (D) for error correction.
- BPs with IBCs in droplets are processed/sequenced as in regular Drop-Seq systems.
- Known barcode sequences are used to esteblish links between cytometric and Drop-Seq data.

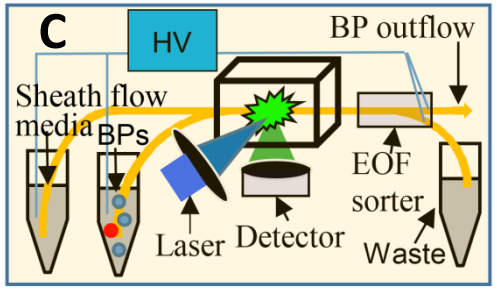


Cytometry and BPs outflow

- Most cytometers could directly supply BPs for pairing only at expense of IBCs overuse because BPs in outflow are diluted with sheath fluid. On-column detection (B) or augmentation of a sheath-flow cytometer with a BP accumulator (A_1) would solve the problem.
- Filed-driven, e.g. electroosmotic systems (C)* with a feedback loop could supply BP's on demand, nearly eliminating IBC waste.
- Droplets-based sorters would require serial, ordered collection of droplets and benefit from the accumulator, but on-column sorters (e.g. C) has no limitations.



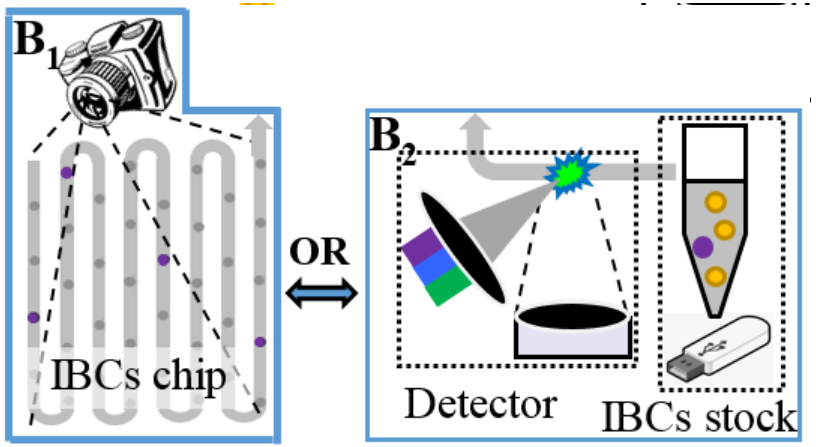




* Andreyev, D. and Arriaga, E.*Simultaneous laser-induced fluorescence and scattering detection of individual particles separated by capillary electrophoresis.* Anal Chem, 2007. **79**(14): p. 5474-8.

IBC fabrication

- Two types of IBC identification: (i) by a serial number in the consecutive outflow from a microfluidic channel (B_1), where serial number is additionally verified by positions of reference beads (B_1 violet) or (ii) by on-flow detection of cores with intrinsic, unique ID (B_2).

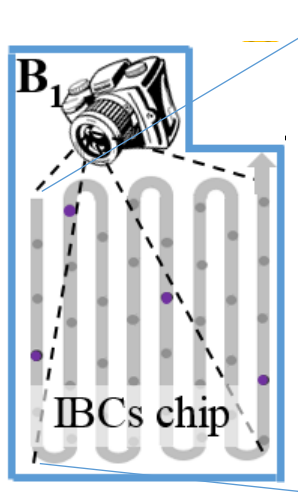


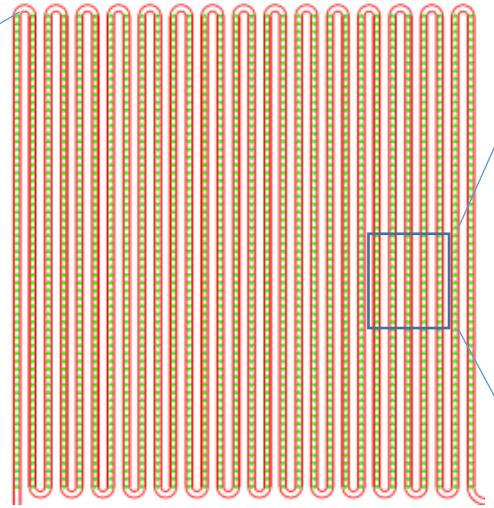
- Synthesis of barcoding oligonucleotides, with defined sequence for each ID, performed via
  - Type i: lithography (similar to “gene chips”) or
  - Type ii: tracked mixing synthesis,

correspondingly.

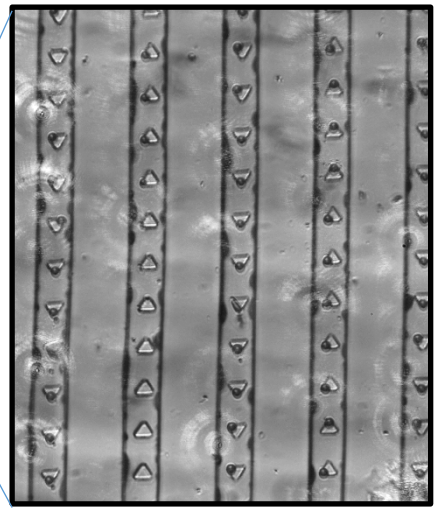
IBC Type I – magnetic immobilization and barcode synthesis

- Cores reversibly immobilized in a microfluidic channel on surface of a chip at pre-defined spots via magnetic, hydrodynamic, or dielectrophoretic traps.
- Oligonucleotides with pre-defined, unique sequence can be grown on each bead-containing spot of the chip, as employed in a “gene chip” fabrication
- After synthesis, beads could be released form surface of the chip into the microfluidic channel and serially eluted.

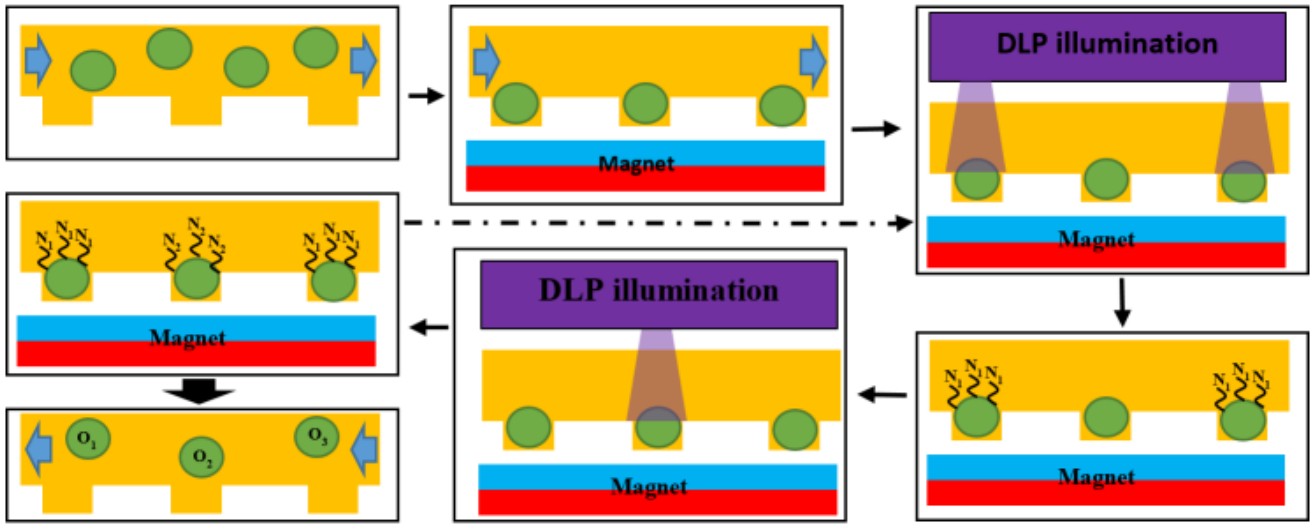




Test chip layout, 3x3mm, 1800 magnetic traps



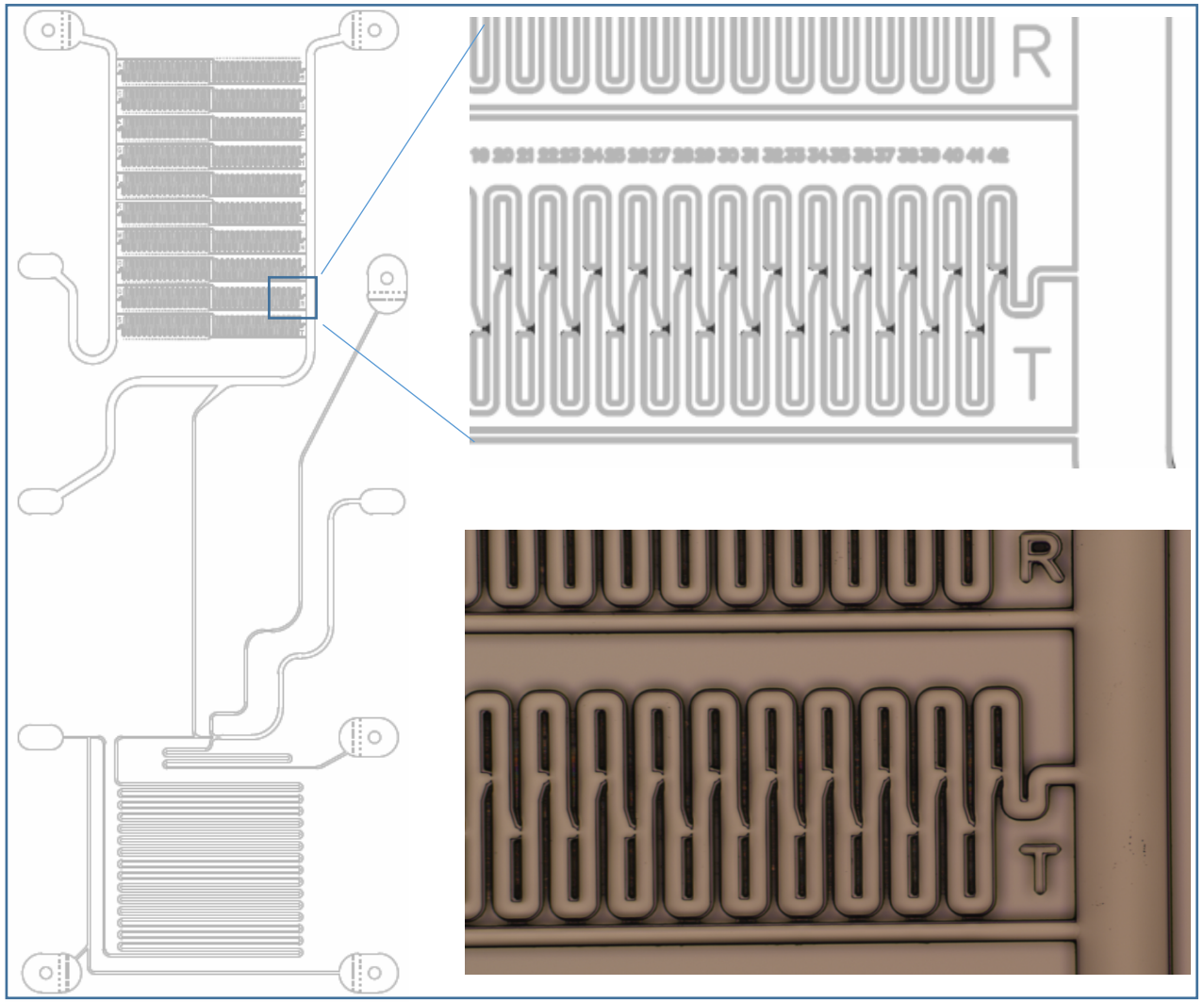
Magnetic IBC cores immobilized on chip



Barcode synthesis on trapped cores using selective photodeprotection

IBC Type I – other immobilization approaches

- Immobilization: hydrodynamic*, dielectrophoretic, optical, etc.
- More sophisticated comparing to magnetic.
- Enable selective, one-by-one core release.

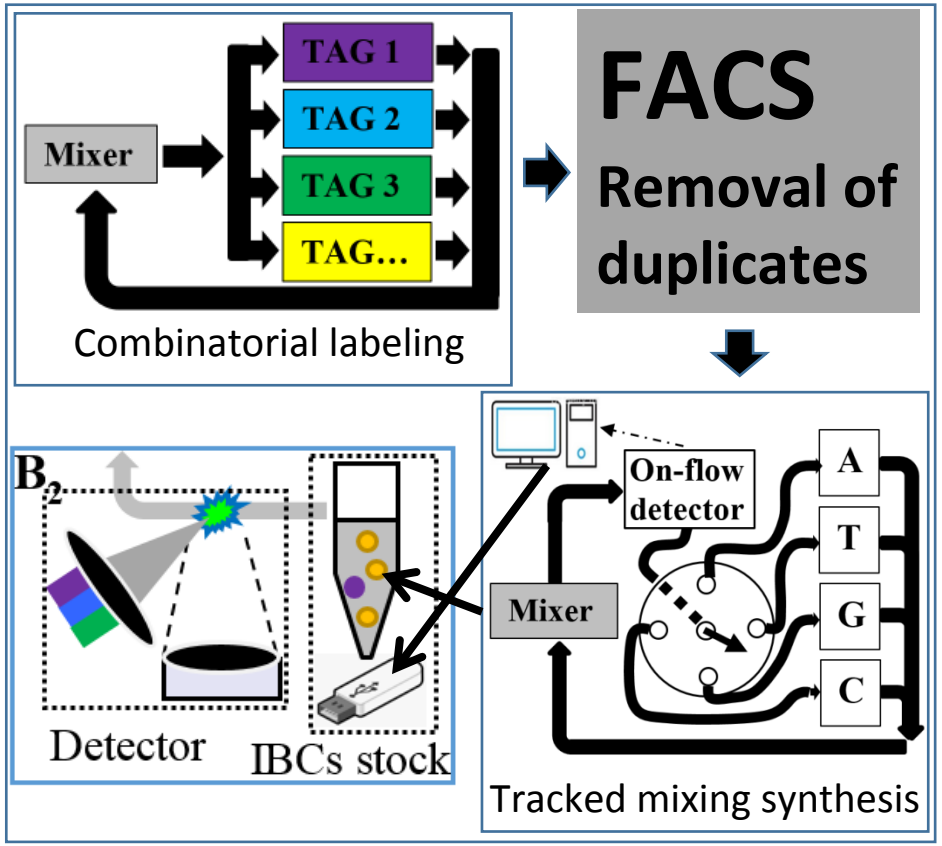


Hydrodynamic traps redesigned from silicon to plastic: a chip layout and image of SU-8 chip/master.

* Kimmerling, R.J., et al., *A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages.* Nat Commun, 2016. **7**: p. 10220.

IBC Type 2 – combinatorial labeling and tracked mixing synthesis

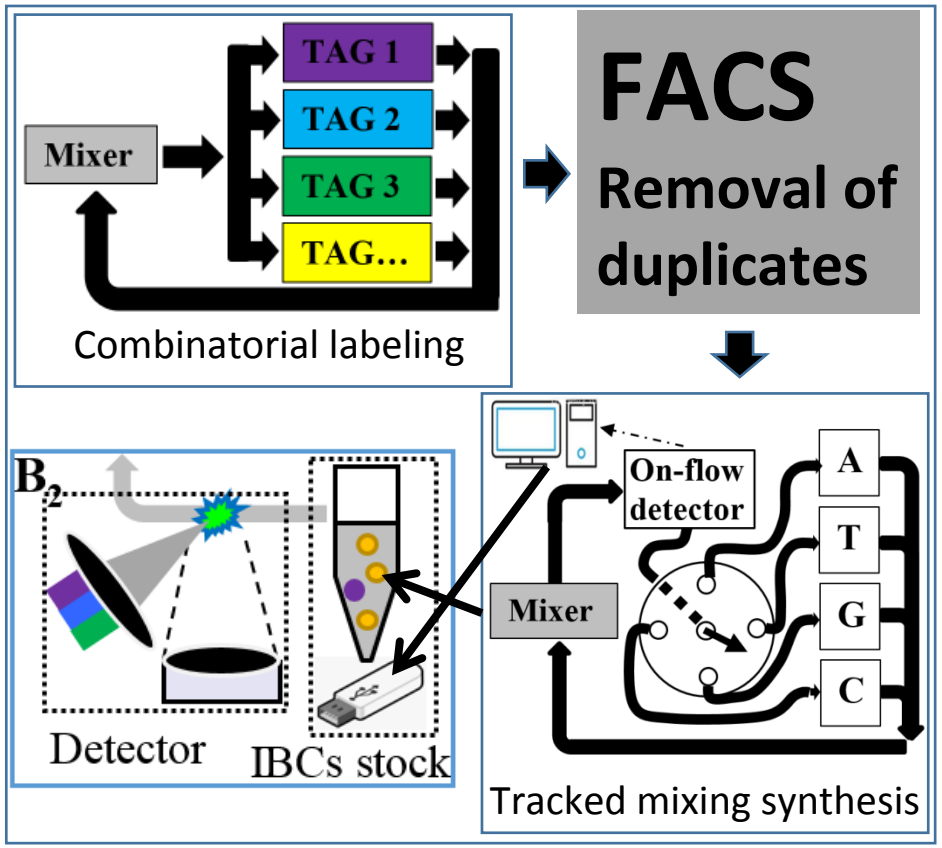
- Cores labeled by combinations of on-flow detectable tags, e.g. by split-and-pool.
- Fluorescent tags are the simplest, and IDs are detectable by any cytometer hardware. Additional tags are possible, including RFID, SERS, magnetic.
- Cores with two fluorophores at ten concentrations, provide 2^{10}=1024 unique IDs, and sold by FlexSet, Luminex.



- 4 fluorophores provide 4^{10}= 1048576 unique IDs. 8 fluorophores for eight -channel cytometer - provide 8^{10}=1 073 741 824, and so on.

IBC Type 2 – combinatorial labeling and tracked mixing synthesis

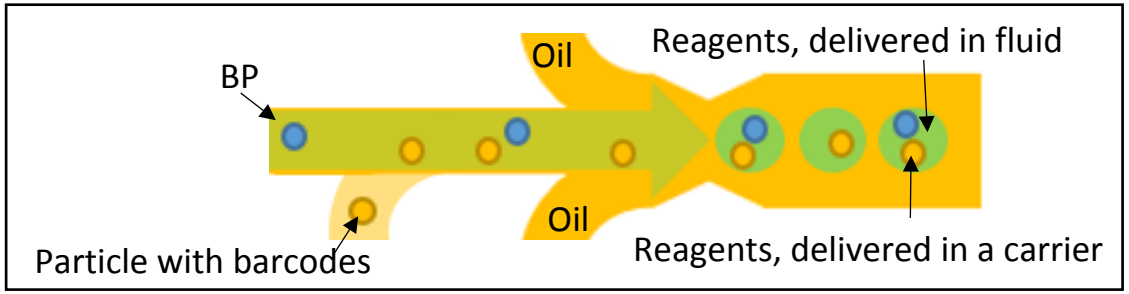
- Cores labeled by on-flow detectable IDs.
- Duplicates with the same ID could be removed using any FACS hardware (software modifications required).
- Tracked mixing synthesis: a core from a mixer is randomly directed to one of four chambers where one of nucleotides is added to it's barcode, and on-flow detector records which nucleotide was added. Next, the core returns to the mixer and cycle repeats.

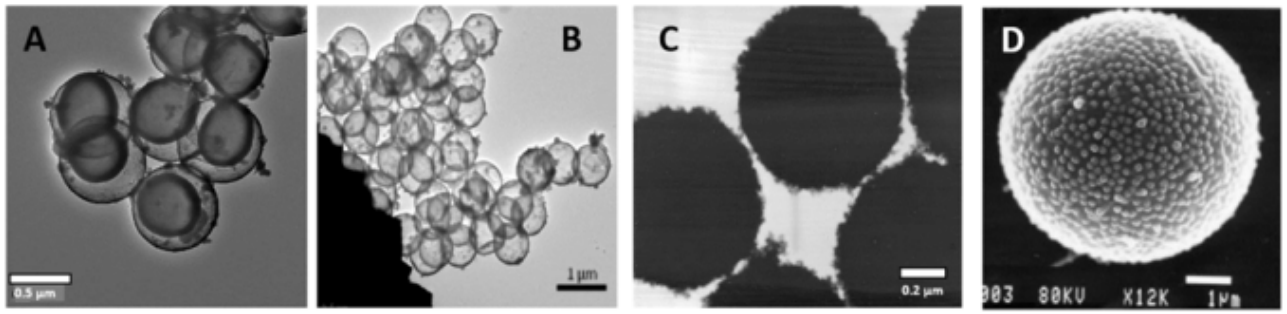


- Resulting set of IBSs include particles themselves and a database with ID -barcode links. Use of such set require separate ID-reader, e.g. additional cytometer.

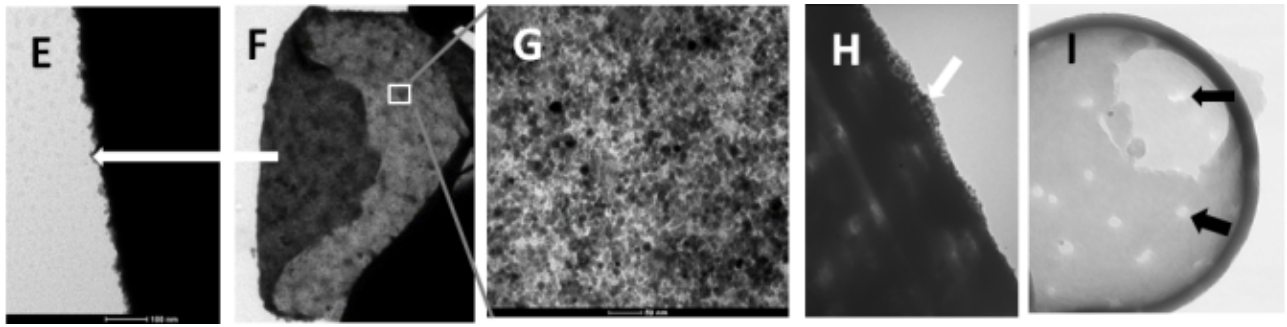
Core engineering

- Existing Drop-Seq approaches deliver reagents via either fluid or barcode carriers, e.g. gel particles.
- While use gel particles for deterministic barcoding is possible, more robust carriers such as hollow silica shells with regulated permeability* could have advantage in delivery, tagging and storage.
- Such shells are modifiable by layers of optical tags and magnetic material, separated by silica to mitigate quenching.
- Other cores, including ones with RFID, SERS, or micro shells produced by aglae can be modified the same way.





A,B: silica shells with partially and completely dissolved templates. C: magnetized template. D: a core with SERS substrate.



E: magnetized, silica encapsulated template. F,G: SiO_2/Fe_3O_4 layer after template dissolution, superparamagnetic domains shown. H: magnetized diatomic thecae. I: SiO_2 shell with custom channels.

* Andreyev, D. , Arriaga, E. *Fabrication of perforated sub-micron silica shells.* Scr Mater, 2007. **57**(10): p. 957-959.

Synchronization and error correction.

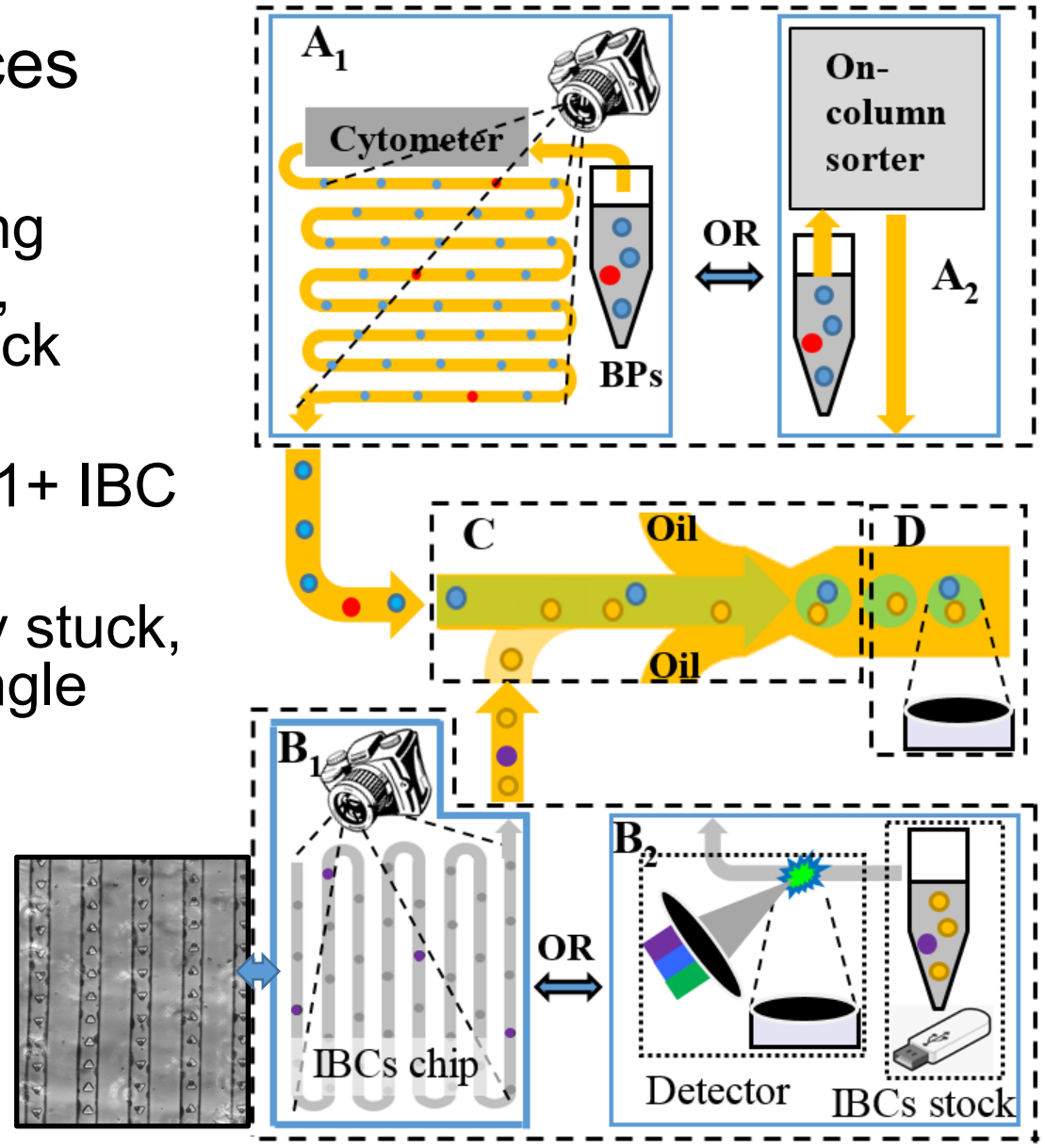
- Linking a database of barcode sequences with cytometric data is prone to errors.

  - For chip-based IBCs, spots for barcoding oligonucleotide synthesis could be empty (B_1), occupied by multiple beads, or beads may stuck outside the spots.

  - For in-drop pairing, the ratio of 1 BP to 1+ IBC per drop might be compromised (D).

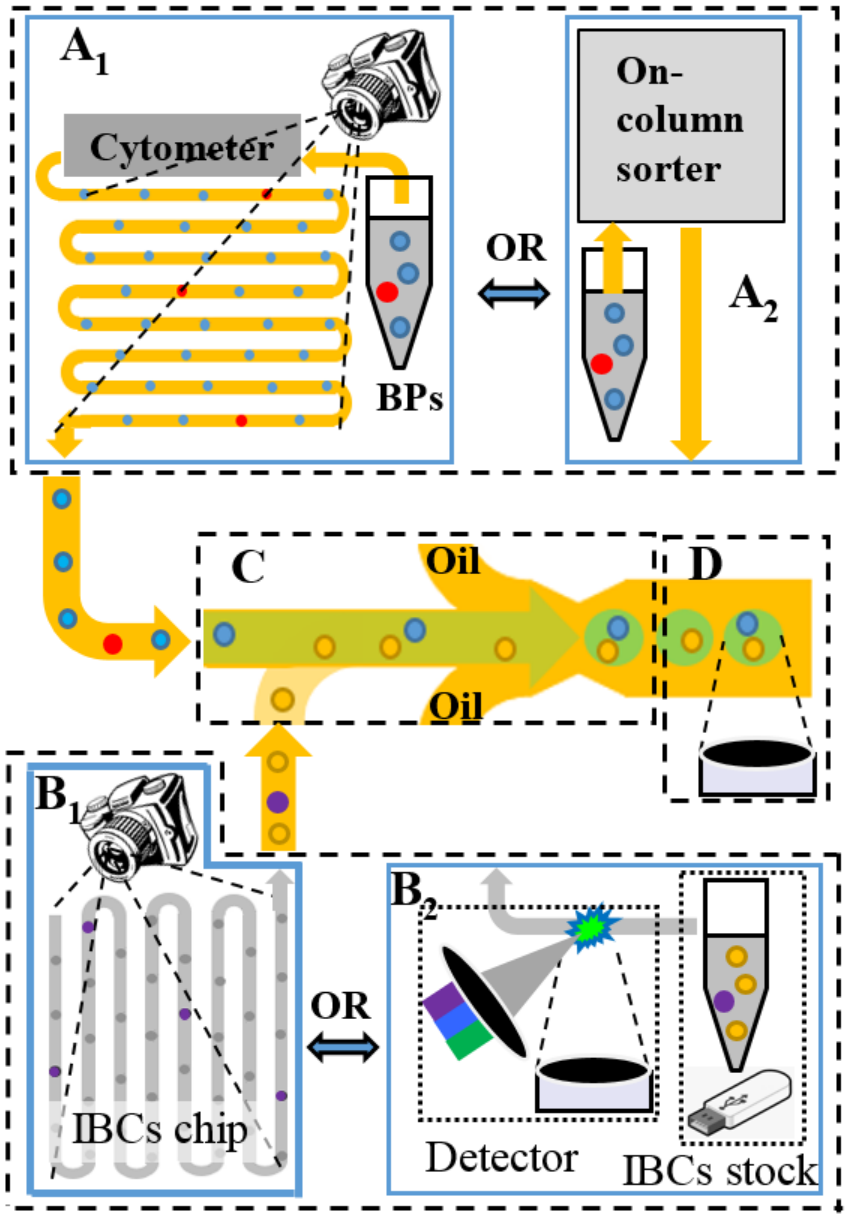
  - After detection in a cytometer, BPs may stuck, switch order, or pairs can be detected as a single event.

- These errors will manifest themselves in <u>mismatching indexes between barcode database and cytometric data</u>, and have to be corrected.



Synchronization and error correction.

- **Pre-sequencing correction:** fraction of fluorescently marked reference beads (A-red B-violet circles) added into the IBCs and BPs, and monitored at cytometer detector (A_2) at IBC on-flow detector (B_2), at droplet monitoring (D), at image of BP accumulator (A_1) and IBC chip (B_1).
- **Post-sequencing correction:** any mismatches between expected by ID and actual barcode sequence is a sign of an error.
- **Repair or delete:** having sufficient information from reference beads and ID-barcode mismatches, index deletions, interchanges, duplications could be repaired. Else, compromised data should be eliminated, such as data for BPs/IBCs between reference beads pair.
- **Applied in parallel**, these methods are expected to provide robust and reliable error correction.



Conclusion

- Integration of Flow Cytometry/FACS (and relatives) and Drop-Seq (and relatives) is essential for life science, as it allows simultaneous analysis of phenotype, functionality, genome and transcriptome of each particle in population of any size.
- There is no fundamental barriers in such integration via deterministic barcoding, only well-established technologies are employed
- Technical difficulties are numerous but fully manageable

- Collaboration is encouraged – we are small and the task is tremendous.

- More discussion on poster session.