Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

Lukas Nögel (M.Sc.)

Born in Breisach am Rhein, Germany

Oral-examination: 31st of January 2019

Multiplexed combinatorial drug screening using droplet-based microfluidics

Referees:

Dr. Martin Jechlinger

Prof. Dr. Benedikt Brors

Summary

The therapy of most cancers has greatly benefited from the use of targeted drugs. However, their effects are often short-lived since many tumors develop resistance against these drugs. Resistance of tumor cells against drugs can be adaptive or acquired and is often caused by genetic or non-genetic heterogeneity between tumor cells. A potential solution to overcome drug resistance is the use of drug combinations addressing multiple targets at once.

Finding potent drug combinations against heterogeneous tumors is challenging. One reason is the high number of possible combinations. Another reason is the possibility of inter-patient heterogeneity in drug responses, making patient tailored treatments necessary. These require screens on patient material, which would drastically benefit from miniaturization, as it is the case in droplet-based microfluidics. However, drug screens in droplets against primary tumor cells have so far only been performed at a modest chemical complexity (55 treatment conditions) and with low content readouts.

In this thesis we aimed at developing a droplet-based microfluidic workflow that allows the generation of high numbers of drug combinations in picolitre-sized droplets and their multiplexed analysis. To this end, we have established a pipeline to produce up to 420 drug combinations in droplets. We were able to significantly increase the number of possible combinations by building a microfluidic setup that comprises valve and micro-titer plate based injection of drugs into microfluidic devices for droplet generation

Furthermore, we integrated a DNA-based barcoding approach to encode each treatment condition, enabling their multiplexed analyses since all droplets can be stored and processed together, which highly increases the throughput. With the established approach we can perform barcoding of each cells' transcriptome according to the drugs it was exposed to in the droplet. Thereby, the effects of drug combinations on gene expression can be studied in a highly multiplexed way using RNA-Sequencing.

We applied the developed approach to run combinatorial drug screens in droplets and analysed the effects of in total 630 drug combinations on gene expression in

K562 cells. The low number of cells needed (max. 2 million cells) for such screens, could enable their application directly on tumor biopsies, thus paving the way for personalized therapy approaches. Since the established workflow is compatible with single cell readouts, we also envision its application to analyse drug resistances in heterogeneous tumor samples on the single cell level.

Zusammenfassung

Die meisten Krebsarten haben stark von der Entwicklung zielgerichteter Medikamente zur Behandlung von Krebs profitiert. Jedoch sind diese oft kurzlebig da viele Tumoren Resistenzen gegen die eingesetzten Medikamente entwickeln. Tumorzellen können adaptive oder erworbene Resistenzen aufweisen, welche häufig durch genetische oder nicht-genetische Heterogenität verursacht wird. Eine mögliche Lösung ist die Behandlung von Krebs mit Medikamentenkombination da mit diesen mehrere Ziele in Tumorzellen gleichzeitig angegriffen werden.

Das Finden potenter Medikamentenkombinationen ist jedoch eine große Herausforderung, da die Anzahl möglicher Kombination immens ist. Zudem kann die Heterogenität zwischen Patienten eine Behandlung zugeschnitten auf individuelle Patienten notwendig machen und somit müssten Medikamente auf Tumorzellen des entsprechenden Patienten getestet werden. Dies kann durch miniaturisierte Medikamentenscreens unter der Verwendung von Tröpfchen-basierter Mikrofluidik ermöglicht werden. Bisher gelang es jedoch nur den Effekt einer geringen Anzahl verschiedener Medikamentenkombinationen (55) in größeren Tropfen (mehrere Nanoliter) auf primäre Tumorzellen zu testen.

Das Ziel dieser Dissertation war es einen Methode der Tröpfchen basierten Mikrofluidik zu entwickeln, welche es ermöglicht eine hohe Anzahl an Kombinationen in Picoliter großen Tröpfchen zu erzeugen. Zu diesem Zwecke, haben wir eine Methode entwickelt, die es ermöglicht 420 Kombinationen von Medikamenten in Tröpfchen zu generieren. Wir konnten eine signifikante Erhöhung möglicher Kombination erzielen, indem wir eine Ventil und Mikrotiter-Platten basierte Injektion von Medikamenten zur Tröpfchen-Erzeugung entwickelt haben.

Zudem wurde eine DNA-basiertes Barcodeverfahren entwickelt, welches es ermöglicht alle Kombination gemeinsam zu prozessieren und analysieren wodurch der Durchsatz von Screens erheblich gesteigert werden kann. Des weiteren wurde das Barcodeverfahren dazu verwendet das Transkriptome der Zellen entsprechend der Medikamentenkombinationen zu kennzeichnen. Hierdurch kann der Effekt welche Medikamente auf die Expression von Genen in Tumorzellen haben, anhand von RNA Sequenzierung untersucht werden. Wir haben das entwickelte Verfahren

angewendet, um die Effekte von insgesamt 630 Medikamentenkombinationen auf die Gene-Expression in K562 Zellen zu untersuchen. Die geringe Anzahl an Zellen (max. 2 Millionen), die für solche Experimente nötig sind, können dazu beitragen dass diese zukünftig direkt mit Tumor-Biopsien durchgeführt werden. Da das Verfahren es zudem ermöglicht einzelne Zellen zu untersuchen, könnte dieses zukünftig dazu verwendet werden die Effekte von Medikamenten auf heterogene Tumorproben auf der Ebene von einzelnen Zellen zu analysieren.

Acknowledgements

This dissertation is the work of four years, during which I met many great scientists and personalities. I'm very thankful for all the inspiration and support I got from these people during the last years.

First of all I want to thank Dr. Christoph Merten for giving me the opportunity to carry out my research in his group at the EMBL in Heidelberg. His supervision and ideas provided the basis of this thesis. I'm thankful for his guidance, feedback and all the discussion we had during my time and EMBL. A special thanks goes to Dr. Martin Jechlinger, Dr. Wolfgang Huber and Prof. Dr. Michael Boutros for being part of my Thesis Advisory Committee. Thanks for your valuable feedback and several stages of my project. Furthermore, I want to thank all examiners of my thesis Prof. Dr. Benedikt Brors, Prof. Dr. Stefan Wiemann, Dr Martin Jechlinger and Dr. Athanasios Typas for supporting my scientific progress.

The input of all former and current members of the Merten group to this project was very helpful. Therefore, I want to thank all of you for your great support during various experiments and for all your feedback during our group meetings. A spacial thanks goes to Dr. Dharanija Madhavan for all the scientific and personal support during the three years we shared an office. I also want to thank Ramesh Utharala for his great support with the Braille display and providing me with LabVIEW software.

The EMBL genomic core facility was of great help in realizing this project and therefore I also want to thank all current and former members. Especially, Dr. Vladimir Benes for helping out with all his experience or a glass of good wine. Additionally, I want to thank my collaborator Dr. Bence Szalai, who analysed the RNA-Seq datasets from the drug screens and provided valuable feedback on the project.

My time at EMBL has been a great experience thanks to all the friends I made during this time. I want to thank all of you for your contribution to my personal and scientific development.

And of course I want to mention Aastha Mathur. Thank you very much for all the great moments during the last years.

And finally I want to thank my family: Mom, Dad, Nele, Deborah and Justus.

Table of Contents

I N T R O D U C T I O N

1. Introduction

1.1. Microfluidics

The term microfluidics refers to approaches, in which small amounts of fluids are handled and analysed in structures with defined geometries (i.e. channels) at the scale of micrometres. The fluid volumes handled in microfluidic devices range between 10^{-6} and 10^{-18} litres depending on the dimensions of used channels, which range between ten and several hundreds of micrometres (Whitesides, 2006). Microfluidic principles have been applied to systems used in biology long before microfluidics as a defined technology was widely used in the field of life sciences. In fluorescence assisted cell sorting (FACS) a sheath fluid is used to focus a stream of cells into a laser beam, enabling high-throughput analysis of single cells (Hulett et al., 1969). Another example is the use of capillary-array electrophoresis for Sanger sequencing, where the use of capillaries with an inner diameter of 50 µm dramatically increased the sequencing throughput during the Human Genome Project (Dovichi and Zhang, 2000). Here a reduction in the fluid volumes used for the analysis of samples caused an increase in throughput. The use of sub-microliter volumes has the advantages of reduced sample consumption, increased sensitivity due to lower diluting volumes and increased throughput. The transition of microfluidics from capillary-based systems towards more sophisticated devices has strongly benefited from manufacturing approaches developed in the field of semi-conductors: The establishment of lithographic methods allowed to produce miniaturized printed circuit boards transporting electrons, and these methods were later on adapted to also enable the manufacturing of microfluidic channels transporting molecules and cells on silicon-based microchips. Fabrications of nano- and microstructures for micro-electronics based on silicon using photolithography was adapted by microfluidics to produce chips in which liquids are directed

through defined geometries (Whitesides, 2006). Although in some early cases microfluidic was done in silicon channels, the development of soft photolithography made microfluidics accessible for a larger research community and provided the basis for its applicability in biology. Here, photolithography is used to produce master moulds with a positive relief of the channel structures. Replicas from master moulds are fabricated by pouring Poly-Di-Methyl-Siloxane (PDMS) over the structures and the cured PDMS casts are removed from the moulds and channels are sealed by covalently bonding it to glass or PDMS (Duffy et al., 1998). This allows rapid and cheap fabrication of microfluidic devices, in which fluid flow is in most cases controlled using external pumps connected with devices over inlets (Eicher and Merten, 2011). A detailed description of microfluidic device fabrication techniques and operations can be found in the Material and Methods section of this thesis. Besides the fabrication, PDMS has the advantage of being biocompatible due to its high gas-permeability and thus allowing long term culturing and imaging of cells (Huberts et al., 2013; Luni et al., 2016), tissues and organisms (Choudhury et al., 2012; Sivagnanam and Gijs, 2013) in microstructures using standard light microscopy.

The injection of fluids into channels of a microfluidic device allows precise control of transported reagents or cells, since flow velocities can be adapted by changing the injected volume over time and, due to the low channel dimensions, turbulent flow is avoided. Instead flow in a microfluidic channel follows laminar flow regimes, since at the generally used scales viscous forces become dominant. This phenomenon can be explained using the Reynolds number (Re), which gives ratio between inertial and viscous forces

$$Re = \frac{\rho v D_h}{\mu}$$

where ρ is the fluid density, v velocity of the fluid and μ the fluid viscosity. Since D_h is defined as the length scale of the system, which in case of microfluidic devices is in the range of micrometres, Reynolds numbers for such devices are generally smaller than 5 (Vyawahare et al., 2010). Therefore, injected water-based fluids behave like viscous fluids and form stable laminar flow regimes in a microfluidic channel. Two or more miscible fluids flowing in parallel through a channel do not mix with each other apart from diffusion, which depends on the Péclet number

(Pe). Pe is given by the channel length l, the fluid velocity v and diffusion coefficient D.

$$Pe = \frac{lv}{D}$$

By adapting channel dimensions and flow rates used for injection, this relationship can be exploited to generate complex concentration gradients within microfluidic channels. How the described and further principles of microfluidics are applied in various systems will be illustrated in the following chapters.

1.2. Single-phase microfluidics

In its simplest form a single-phase microfluidic system can be a chamber providing a micro-scale environment, with precise temporal control of its composition by adaptations of the injected reagents. Yet these setups provide a powerful tool to analyse and perturb complex biological system as recently described by studying signal dynamics during mouse embryonic development (Sonnen et al., 2018). Embryonic tail buds were placed in a microfluidic chamber, where the effects of temporally controlled drug pulses were studied using long-term imaging to decipher the relationship between Wnt and Notch signalling during segment formation. Due to low dimensions, such a chamber can be easily modified to obtain laminar flow of two or more reagents by their continuous injections. This allows the generation of different environments in a single chamber and thereby to study their spatial and temporal effects on the development of organisms (Lucchetta et al., 2005).

Furthermore, laminar flow in micro-channels can be used to generate concentration gradients by adapting the channel length and flow rates to obtain sufficient mixing by diffusion (Péclet Number < 1) as fluids flow through the channel (Kim et al., 2010). Gradients of chemokines find wide applications in microfluidics to analyse their effects on cell migration. Rapid established chemokine gradients (2 min) were used to study effects of CCL21 and CCL19 on dendritic cell migration and found different responses based gradient concentrations (Haessler et al., 2011). This study illustrates how fluid dynamics in micro-channels can be applied to complex biological systems and provides the basis for a more systematic and quantitative analysis.

While the number of distinct environments generated by laminar flow is limited, compartmentalization of microfluidic devices by valves are used to increase the diversity of conditions and reactions on a single device. Most widely used are Quake valves, which make use of the elastic nature of PDMS to open and close channels by pneumatic pressure (Unger et al., 2000). In these more sophisticated designs, channels can be closed to facilitate compartments disconnected from other regions of the chip. Most prominently, these valves are used for single cell sequencing of genomes (Fan et al., 2011; Wang et al., 2012) and transcriptomes in the Fluidigm C1 system by trapping and barcoding of single cells in chambers. However, valve-based compartmentalization has also proven its power for other systems genomic approaches such as identification of binding motifs for a large set of transcription factors (Isakova et al., 2017). Furthermore, confining single cells in chambers using valves increases the sensitivity for various proteomic approaches since the dilution factor of proteins released from a cell can be reduced dramatically. This enabled the simultaneous detection of secreted proteins from thousands of single T-cells by antibodies and could identify functional heterogeneity between cytotoxic T-cells isolated from cancer patients (Ma et al., 2011). Since convectional secretion assays (e.g. ELISA) cannot detect several secreted proteins at once and furthermore lack the sensitivity to distinguish between the numbers of secreted proteins down to several hundreds, these findings would have been impossible to achieve in conventional tissue culture formats such as micro-titre plates.

1.3. Two-phase droplet-based microfluidics

Valve-based compartmentalization in a one-phase system is still limited to several thousand chambers per device. Using two-phase systems, this limitation can be overcome by several orders of magnitude. Two immiscible reagents are used to generate discrete units from one reagent by a continuous phase of the other reagent. For this work, the two phases refer to a continuous oil phase, which is used to form droplets from an aqueous phase. Channel geometries are used to produce droplets from an aqueous phase with defined size and content. Depending on the geometries of the channels and the flow rates, the volumes of produced droplets are between 50 fl and 1 nl. Since each of these droplets form a confined compartment, each droplet is considered as a reaction vessel, which can be produced in a passive process at kilohertz frequencies (kHz). The basic concept of

droplet production is to form an interface between two immiscible fluids (e.g. water and oil), followed by the passive segregation of one phase (water) into the continuous phase (oil). Early reports used a co-flow system in which the tip of a capillary is used to inject one phase into a stream of an immiscible phase. The stream of the continuous phase breaks the aqueous phase into droplets (Umbanhowar et al., 2000). Then Quake and colleagues introduced the concept of a T-junctions for droplet production. Here, droplets were formed by two immiscible phases being injected orthogonal to each other. The continuous phase shears off droplets at the T-junction, whose size can be adapted by changing the ratio of the flow velocities (Thorsen et al., 2001). A third approach for droplet production was introduced, which makes use of a flow-focussing junction. The formation of droplets at a flow-focussing junction is achieved by the concentric acceleration of the dispersed phase by the continuous phase followed by a constriction that opens into a wider channel. The acceleration of an aqueous phase by a continuous oil phase, results in a narrow aqueous stream that breaks into droplets due to the constriction (Anna et al., 2003). In order to form a stable emulsion for long-term storage, the surface tension between the immiscible phases needs to be reduced, which is achieved by adding surfactant to the continuous phase. Since fluorinated oils are widely used as a continuous phase for droplet production, surfactants made of perfluorinated poly-ethers linked to polyethylene-glycol were developed (Holtze et al., 2008). The popularity of fluorocarbon oils and surfactants arises from its high biocompatibility (Giaever and Keese, 1983), its high oxygen and carbon dioxide solubility (Lowe et al., 1998) and its low solubility of hydrocarbons (Lonostro, 1995). These properties allow the encapsulation of cells into droplets and their incubation in droplets over several days (Clausell-Tormos et al., 2008).

1.3.1. Droplet manipulations

The small volume but high numbers of samples used in droplet-based experiments, made it necessary to develop specialized tools to perform automated manipulations of droplets after they have been generated. After off-chip incubations, droplets can be reinjected into microfluidic devices for their analysis and manipulations. Droplet analysis often includes fluorescence-based measurements for the quantification of biochemical assays after an incubation time. For this purpose a laser beam is focussed onto the channel network and emitted light is quantified by

photomultiplier tubes (PMTs). Due to the low timescale of fluorescence detection (sub-milliseconds), it is well suited for high-throughput droplet-based microfluidics (Solvas et al., 2011). Droplet manipulation steps include processes that aim at mimicking liquid handling steps performed at macro-scale.

Droplet fusion

Merging of droplets is used to add reagents to start or stop reactions within a droplet or for dilution of droplet content. This operation is comparable to a pipetting step used for marco-scale liquid handling. Merging two droplets can be achieved in a passive process in which two droplets are brought into contact while their interfaces are destabilized by an expansion of the channel width (Bremond et al., 2008). Pairing and fusing two droplets was also demonstrated at kHz frequencies by injecting surfactant stabilized droplets into a channel and pairing each of them with a non surfactant-stabilized droplet (Mazutis et al., 2009). An active way of merging two droplets is their contact-based fusion under an electrical field generated by electrodes along the channel (Priest et al., 2006). These methods are powerful for running sequential and complex reactions within the same droplets like the fragmentation of DNA followed by its amplification (Lan et al., 2016).

Similar to droplet fusion, reagents can be added to pre-existing droplets in a process called pico-injection. Surfactant stabilized droplets are injected into a microfluidic device and spaced out using oil. Droplets pass through a narrow channel and pass by a pico-injector nozzle, which is used to inject reagents. An electric field generated by an electrode opposite of the nozzle destabilizes the surfactant film surrounding droplets and allows reagents from the nozzle to be injected into each droplet passing by (Abate et al., 2010). More details can be found in the material and methods chapter of this thesis.

Droplet Sorting

Sorting of droplets is performed to enrich specific droplets based on their properties. Fluorescence activated droplet sorting (FADS) is considered the droplet based counterpart to FACS and does include the detection of a fluorescence reporter followed by a sorting decision that results in the collection or disposal of a droplet. The main advantage compared to FACS based sorting is that instead of

sorting cells, entire assays (including soluble reagents) are sorted and thereby rare screening results can be enriched from a large population (Baret et al., 2009). The basic concept of sorting droplets is their detection upstream of a channel bifurcation and based on the measured signal, the trajectory of the droplet is either changed to direct droplets into the collection channel or droplets are send to the waste by not interfering with the streamline. By applying an electric field when a droplet of interest passed the detection point, the droplet is deflected into the collection channel (Baret et al., 2009). Such methods were used for directed evolution of enzymes encapsulated along with fluorescence reporter substrates. Iterative selections of horseradish peroxidase by FADS were performed to select enzymes with increased catalytic activity from a library with 10^8 mutants in about 10 hours which is a thousand fold increase in speed compared to conventional plate based screens (Agresti et al., 2010).

1.3.2. Cells in droplets

As already discussed before, the compartmentalization of cells into small volumes is advantageous for many assays since the analytic performance can be enhanced. Encapsulating cells into droplets is therefore of great interest since it allows rapid production of cell containing compartments. By injection of a cell suspension into a device for droplet-generation, the segregation of the suspension into droplets results in the distribution of cells into discrete units. The number of cells in each droplet is dependent on the cell concentration and the size of the droplet. Since the droplet size is constant at a given flow rate, the number of cells per droplet can be controlled by adjusting the density of cells, resulting in a Poisson distribution of probabilities for the number of cells per droplet. For cell concentrations smaller than 10^6 cells/ml and 660 pl sized droplets, the probability of empty droplets is the highest (~0.6), followed by a probability of ~0.3 for having one cell in a droplet (Clausell-Tormos et al., 2008). Due to the low probability of obtaining droplets with more than one cell ($p<0.07$) and the high frequency of droplet production, droplet based microfluidics has become a popular tool to study single cells. By encapsulating heterogeneous cell populations into droplets, each droplet containing a single cell represents a genetic or phenotypic unique unit with high local concentrations of metabolites. This advantage was used in a functional antibody screen using hybridoma cells (El Debs et al., 2012). This is a well-suited

example to illustrate the power of droplet-based microfluidics: A by definition heterogeneous cell population was encapsulated at the single cell level to analyse the function of proteins secreted by the cells and select for droplets containing a cell of interest. Translating hybridoma cell based antibody screens into droplet-based microfluidic can accelerate the discovery of new potent monoclonal antibodies. Due to the low dilution of antibodies in picolitre volumes, it is feasible to study antibodies secreted from primary B-cells since femto-gram concentrations are detectable (Eyer et al., 2017; Shembekar et al., 2018). A limitation for screening cells in droplets is long-term cultivation due to depletion of media and the accumulation of toxic metabolites. The cell viability of cells in picolitre-sized droplets generally drops after two or three days (Clausell-Tormos et al., 2008). This problem can be partially compensated by increasing the droplet size. Another possibility is the co-encapsulation of cells with agarose, which polymerizes to form solid spheres. This allows removal of the oil phase and long-term cultivation of cells surrounded by media to form spatially defined spheroids (Sart et al., 2017).

1.4. Drug combinations

The development of new drugs can be very expensive, time consuming and approval of new compounds is difficult. The two major reasons for drugs not getting an approval for the market are low efficacies and or safety concerns (Kola and Landis, 2004). Low efficacies can be caused by redundancies in signalling networks, resulting in the activation of alternative routes to compensate for the inhibition of one protein (Jia et al., 2009). Additionally, preclinical studies on animal models often lack clinical transferability and therefore efficacies for drug candidates can vary between preclinical and clinical phases (Hackam and Redelmeier, 2006). Toxicity of drugs can arise from drug pleiotropy due to non-selective interactions with off-targets (Klaeger et al., 2017) and from diverse functions of drug targets. Drug combinations have the potential to overcome both problems, by increasing efficacies of drug treatments and thereby reducing the toxicity since lower concentrations can be used. Furthermore, drug combinations from pre-existing and approved drugs offer a cheap and fast way to overcome resistances and to define new treatments with enhanced efficacies and therefore play an important role in the development of new treatment strategies.

A drug combination can be defined as good, when the effect of the combination is stronger than the summed effect of the individual drugs, an effect called drug synergy. Synergistic effects are caused by three main actions: Anti-counteractive actions, complementary actions and facilitating actions (Sun et al., 2013). An anti-counteractive actions result from inhibiting the response caused by one drug with a second drug. The disruption of the peptidoglycan cell wall by one drug would eventually activate the cell wall repair machinery. The additional inhibition of the cell wall repair system with a second drug can have a synergistic effect, since it can accelerate bacterial cell death. Drug synergy from complementary actions is induced by drugs that target proteins of the same or different pathways, which are important for the regulation of the same biological functions. Thereby bypass reactions by activation of alternative signalling routes can be prevented. This approach is common for therapies used in cancer where several drugs aim at the induction of apoptosis. In synergy due to facilitating actions one drugs improves the efficacy of another drug. Such synergy was observed for the inhibition of endogenous ligands by one drug resulting in an enhanced binding of a second since the competition for binding the target was reduced (Jia et al., 2009). The increase in drug efficacies often allows treatments at lower concentrations and thereby a reduction in toxicity is achieved. Besides synergism, drug combinations show antagonistic effects caused by interfering or counteractive actions. Treatment with a drug that incorporates into the DNA in combination with a drug that induces cell cycle arrest could have a counteractive effect.

Finding good drug combinations is challenging since the number of combinations increases exponentially with every new drug and thus often exceeds feasibility. All pairwise combination of FDA-approved drugs (~1500) results in over one million combinations. In order to accelerate the process of finding new combinations, high-throughput drug screening approaches and computational methods have been developed. Multiplexed drug screenings, in which compounds are pooled and then screened for synergy, decreases the number of combinations in a first screening round. Pools of drugs with enhanced efficacies were tested in a second screening round to identify active combinations (Borisy et al., 2003; Tan et al., 2012). Tan et al. used this approach to screened $5x10^5$ drug combinations with synergistic effects against HIV replication and used less than 3% of the wells compared to conventional pairwise drug screens. Drug combinations harbour great potential for

cancer treatment as demonstrated by a recent study that predicted and tested combinations from 218 compounds to find synergistic drug pairs on cells from leukaemia patients (He et al., 2018). The use of robotics for large-scale liquid handling on one hand enables to screen high numbers of drug combinations but on the other hand these screens are expensive and often not feasible for biological samples with small cell numbers like tumor biopsies. Computational approaches aim at overcoming these limitations by using training data to predict sensitivities towards drugs. Pharmacogenomics integrates high dimensional datasets from multiple sources (genomics, transcriptomics, proteomics etc.) and drug responses to build models, which associate genotypes and phenotypes with drug sensitivities. These models provide predictions on how a drug response is affected by genotypes or phenotypes and thereby knowledge on multiomics data from patients or disease models can be used to tailor treatments (Bansal et al., 2014; Menden et al., 2018). Crucial for good predictions of drug combinations is the availability of large training datasets and public available resources like databases for pathway activity and drug properties. The lack of medical relevant training datasets (e.g. drug screenings on tumor biopsies) often results in a reduced translatability of computational models. Additionally, the readouts of conventional drug screenings are relative simple (e.g. growth reduction) and do not necessarily account for dynamic responses of drug treatments. Therefore, the development of new strategies for combinatorial drug screenings with deep characterizations of drug response could provide a powerful solution for defining new treatments based on drug combinations.

1.4.1. Drug screening and microfluidics: A perfect match?

In large-scale screens using robotics, a reduction in reagents and material is achieved by increasing the number of wells per micro-titre plates. Droplet-based microfluidics offers a dramatic increase in the number of compartments since volumes drop to picolitre scale, resulting in a massive reduction in sample consumption. Performing drug screens in droplets seems like a great solution, especially for drug combinations. Despite its great potential to enable cheaper and faster screens of drug combinations, microfluidics was and still is lacking behind its promise. This discrepancy is due several limitations that arise when translating drug screens into droplet-based microfluidics. A bottleneck for generating chemical

diverse droplets is that the composition of the aqueous phase injected into a drop-maker needs to be diverse as well. For emulsions with chemically diverse droplets, different compounds must be loaded onto a microfluidic device, which is very slow (minutes) compared to droplet production (up to several kHz) and thus limits the throughput. Stabilizing droplets with surfactant generally occurs above the critical micelle concentration and consequently micelle formation is a common phenomenon in emulsions. Since micellar structures can transport small molecules between droplets, retention of a compound within droplets depends on its chemical properties (Gruner et al., 2016). Efficacies of drug combinations are generally assessed by determining a reduction in growth, which depends on long-term incubations (e.g. 72h). However, at these timescales depletion of growth media can become problematic when using subnanolitre volumes. Additionally, the rate in which cells proliferate in droplets is slower as compared to plate-based systems and limited to suspension cells (Clausell-Tormos et al., 2008). Furthermore, in plate-based drug screens each drug is assigned to a specific well on a plate, allowing its identification and pairing of treatment conditions with assay readouts. Free-floating droplets of an emulsion stored together in a tube have no spatial order. However, information on droplet contents is crucial for the identification of hits. Several studies have demonstrated how these limitations can be overcome or avoided.

A first proof-of-concept study performed drug screens in droplets used a drug library containing Mitomycin C at 8 different concentrations, which was encapsulated into droplets. Each concentration was encoded by a unique concentration of a fluorescence dye enabling their detection. Droplets were fused with droplets containing cells and after an incubation of 24h cell viability and cell death was measured (Brouzes et al., 2009). The power of droplet-based microfluidics for screening compound libraries was shown by Miller and colleagues who made use of dispersion to dilute samples in a tubing. Thereby high-resolution dose-response relationships (10,000 concentrations) between an enzyme and more than 700 compounds were analysed (Miller et al., 2012). Here concentrations were also encoded by the dilution of a fluorescence dye and substrate conversion by the enzyme was measured only a few minutes after the encapsulation and thus drug exchange was negligible. While these two approaches only screened single compounds, the encapsulation of pairwise combinations required more complex

strategies. Generation of drug combinations in droplets was achieved by droplet fusion of two droplets containing a single compound or valve-based approaches where compounds from several inlets were encapsulated into droplets.

To exemplify valve-based approaches, a recently published workflow developed in our laboratory will be described. Valves were applied to produce combinations of drugs in large droplets called plugs (~500 nL) stored sequentially in tubings. The valve system was set up by using the pins of braille display, normally used for blind people to read, to close and open channels. A microfluidic chip bound to an elastic membrane is aligned with its channels on top of the pins and their actuation was used to either direct injected drugs for plug production or the waste outlet. Since the produced plugs were stored in a sequential order using microfluidic tubings, the information on plug composition was maintained. Additionally, plugs were spaced out by mineral oil preventing their coalescence so that no surfactant for their stabilization was necessary preventing surfactant-based drug exchange. In the presented study this system was used to screen tumor biopsies against 45 drug combinations (Eduati et al., 2018). The low cell number necessary in each screen allowed testing drug combinations directly on tumor biopsies. The valve-based approach presented by Eduati et al. was used and developed into the workflow for screening drug combinations in picolitre-sized droplets presented in this thesis.

1.5. Next-generations sequencing and microfluidics

The development of todays widely used sequencing platforms was largely dependent on microfluidics. Emulsion PCR based Roche 454 sequencing or flow chambers for solid-phase based sequencing used by Illumina systems apply microfluidic principles such as reduction in sample volumes to perform high-throughput sequencing (Metzker, 2010). Whereas standard protocols for preparing sequencing libraries does not include microfluidic workflows, the advantages of compartmentalizing reactions during these steps is becoming more popular, especially for single cell sequencing assays. High-throughput sequencing or next-generation sequencing (NGS) of genomes relies on the fragmentation of input material (i.e. Genomes). Millions of DNA-fragments are sequenced and the obtained reads are subsequently aligned to a reference genome. Due to amplification errors between different fragments during library preparation, the coverage over a

genome (number of reads at a certain position) often fluctuates, which causes errors when detecting copy number variations. Emulsifying single DNA fragments into droplets for their amplification reduces amplification biases and results in a uniform coverage of genomes even for single cells (Fu et al., 2015). Furthermore, short-read sequencing for *de novo* sequencing or sequencing of mixed populations is difficult, since short reads need to be assembled into genomes. Barcoding of DNA in droplet offers a potential solution as shown by Lan and colleagues: The encapsulation of long DNA templates (~5 Kb) into droplets for amplification and fragmentation was followed by a barcoding step **(Fig. 1.1A)**. Droplets containing barcoded primers were fused with droplets of fragmented DNA and in an overlap-extension PCR all fragments derived from one template were labelled with one barcode (Lan et al., 2016). By sequencing the barcode and the DNA fragments, all fragments can be assigned to their original template DNA, simplifying their assembly. This method was further developed and applied to barcode genomes from single prokaryotic cells and thus enables single cell genome sequencing of diverse bacteria populations (Lan et al., 2017).

NGS is also widely used as a method to sequence all mRNA molecules present in cells, the so-called transcriptome, by reverse transcribing them into cDNA. Sequenced fragments, or reads, are subsequently aligned to a reference genome. The quantification of expression levels for certain genes allows making statements on the cellular state. Therefore, RNA-Sequencing (RNA-Seq) has become extremely popular for the analysis of transcriptomes from single cells (scRNA-Seq), since heterogeneous cell populations can be analysed at great resolution and throughput. Here, the capacity of droplet-based microfluidics for compartmentalizing single cells has become a powerful tool. Single cell are encapsulated into droplets where each cell's transcriptome is labelled with a unique DNA-barcode **(Fig. 1.1B, C)**. By paired sequencing of barcodes and cDNA fragments, each read is assigned to a barcode and thus to a single cell (Klein et al., 2015; Macosko et al., 2015). These methods (Drop-Seq and inDrop) allow high-throughput sequencing of thousands of single cell due to randomly distributing DNA-barcodes together with cells into droplets

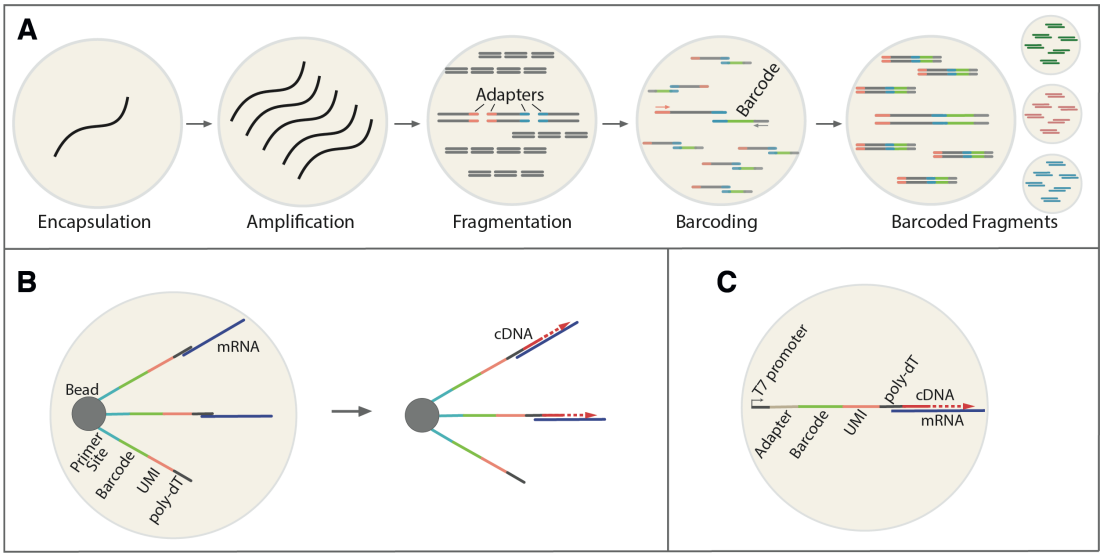

Fig 1.1: Barcoding approaches in droplet-based microfluidics

(A) Barcoding of DNA templates as described by Lan et al.: Single DNA template are encapsulated into droplets together with reagents for their amplification. Droplets containing amplified templates are merged with droplets containing transposases resulting in the fragmentation of DNA templates and the introduction of constant adapters. Each droplet with fragments from one DNA template is merged with a droplet containing a unique DNA barcode and PCR reagents. Upon fusion, a constant part in the barcode sequence hybridizes to a constant adapter and fragments are barcoded by overlap-extension PCR. **(B)** Barcoding of mRNA using Drop-Seq: Single cells and beads coated with barcoded primers are encapsulated into droplets. Cells are lysed and released mRNA hybridizes to poly-dT sequences of the barcoded primers. Droplets are broken to purify beads and mRNA. Reverse transcription of mRNA into cDNA results in barcoding of all mRNA molecules from one cell with a cell specific barcode and a unique molecular identifier (UMI). **(C)** Barcoding of mRNA with inDrop: Single cells are encapsulated together with hydrogel beads containing barcodes. Barcodes are released from the hydrogels and mRNA from lysed cells hybridizes to their poly-dT sequences. Reverse transcription is directly carried out in droplets resulting in the barcoding of mRNA from one cell with a cell specific barcode and UMIs. The T7 promoter allows linear amplification of cDNA by *in-vitro* transcription of barcoded cDNA after they are released from droplets.

1.6. The biology of cancer

Cancer is a disease in which genomic changes result in the malignant transformation of cells. Those cells (called cancer or tumor cells) are defined by uncontrolled cell division and the capability to invade other tissues (Hanahan and Weinberg, 2000). According to the International agency for research on cancer, 14.1 million new cases of cancer were reported worldwide in 2012 and 8.2 million people died from cancer. There are several types of cancer, defined by the tissue they originated from. Lung cancer, prostate cancer, colorectal cancer and stomach cancer have the highest prevalence in males, whereas females suffer most frequently from breast cancer, colorectal cancer, lung cancer and cervical cancer.

1.6.1. Cancer as a genetic disease

Natural genetic variation is crucial for any population since it provides the basis for better adaptation. The fusion of two gametes (haploid genomes) during fertilization gives rise to an embryo with two copies of each gene, called alleles. This process facilitates natural variation since two genomes form a new one in which two alleles can either have the same sequence (homozygosity) or two alleles show genetic variation in their sequence (heterozygosity). Genetic variations in the germlines affect the entire organism and can be passed on to the next generation (Alberts, 2015). In contrast, somatic variations occur outside of the germlines and manifest itself generally only in a small subset of cells. Genomes of cells accumulate mutations (i.e. genetic variations) over their lifespan, which are passed on to daughter cells during cell division. There are different classes of variations, which are defined based on the number of base pairs (bp) they are affecting. Single nucleotide variants (SNVs) are 1 bp substitutions and are the most common form of genetic variations (Genomes Project et al., 2012). Insertion and deletion (called Indels) of nucleotides can also only affect a single nucleotide, however, Indels can be more severe since they can cause frame-shifts resulting in an altered reading frame of a gene. Genomic variations that affect DNA segments larger than 1 kb are called structural variations (SVs). Types of SVs include copy-number variations (deletion, insertions and duplications of DNA segments), inversions (the orientation of DNA segment has changed) and translocations (DNA fragment is integrated at a new position) (Feuk et al., 2006).

Genetic variations can have an impact on cellular functions due to the central dogma of molecular biology: Genetic information stored in genomes is used by cells to synthesize proteins, which is facilitated by the two processes called transcription and translation (Crick, 1958). The transcription of a gene gives rise to mRNA, which nucleotide sequence is translated into an amino acid sequence by decoding always three nucleotides (codons) into one amino acid. Since the AS sequence determines the structure of a protein and thus its function, changes in the genomic information can affect cellular processes carried out by proteins (Alberts, 2015). Non-synonymous substitutions of nucleotides alter the AS-sequence of a protein and thereby proteins with altered functions are produced (Ramensky et al., 2002). Indels with frame-shift mutations often result in premature stop-codons and thus truncated proteins or non-functional proteins are produced from the mutated gene (Lin et al., 2017). Structural variations can have more severe effects on an organism. Copy number variations may cause the loss of proteins encoded by a region (deletions) or the overproduction of a protein due to its excess transcription (duplications). Translocations can alter the regulatory environment of genes and thus cause losses as well as gains in expression (i.e. transcription) of the corresponding genes (Rodriguez-Revenga et al., 2007).

Genomic variations in genes responsible for the regulation of cell growth and survival have the potential to initiate uncontrolled cell proliferation and thus can lead to the development of cancer. Those genes are classically defined as oncogenes or tumor-suppressor genes. An oncogene derives from a proto-oncogene (i.e. a gene with growth promoting functions) by a gain of function mutation, resulting in a constitutively activated or highly abundant protein (Croce, 2008). A well-studied proto-oncogene is *BRAF*, which is part of the mitogen-activated protein kinase (MAPK) pathway that transmits binding of a growth factor into a cellular response (Peyssonnaux and Eychene, 2001). Oncogenic activation of *BRAF* is caused by point mutations that result in elevated kinase activities and growth factor independent activation of MAPK pathway by BRAF (Davies et al., 2002). Proteins derived from tumor-suppressor genes are involved in the regulation of cell cycle and apoptosis and balance growth promoting proto-oncogenes. A loss of function mutation in a tumor-suppressor gene (generally both alleles must be affected), promotes uncontrolled growth of a cell (Weinberg, 1991). The *TP53* gene is mutated in around 50% of all cancer due to its central role in regulation of DNA repair, cell

senescence and apoptosis (Alberts, 2015). More recently the concept of driver mutations and passenger mutation has been introduced. Driver mutations are those that promote tumor development and progression, whereas passenger mutations are often the result of genomic instability and have no or only little impact on tumor progression (Pon and Marra, 2015).

1.6.2. Hallmarks of cancer

For the transformation from a neoplastic lesion with increased proliferative activity into a malignant cancer, in most cases a gradual accumulation of mutations is required. The presence of a single driver mutation does not result in the development of cancer. Multiple events are necessary for cells to overcome constraints of cell growth and acquire the potential to invade and colonize distant tissues. Although different cancer types show great diversity, there are is a set of properties cells need to acquire for their transformation, which are common for all types. These properties are called the hallmark of cancer (Hanahan and Weinberg, 2011):

Sustained Proliferative Signalling: Induction of cell division independent on growth factors

Evading growth suppression: Liberation from mechanisms that negatively regulate cell growth

Resisting cell death: Disruptions of mechanisms inducing apoptosis in cells

Enabling replicative immortality: Unlimited proliferation by avoidance of telelomere-shortening induced cell senescence

Inducing Angiogenesis: Ensure supply of nutrients and oxygen for excessive cell growth by blood vessel formation

Activating invasion and metastasis: Escape of cells from the primary tumor and colonization of distant tissues

Genomic Instability: Enables other hallmarks by an accelerated acquisition of mutations

Chronic inflammation: Enables acquisition of hallmark characteristics by a providing a mutagenic and proliferative environment

Reprograming energy metabolism: Glycolysis-dependent energy production under aerobic and anaerobic conditions

Evading immune destruction: An immune suppressive microenvironment is established to avoid immune-based elimination

1.6.3. Tumor heterogeneity

The acquisition of hallmark characteristics is a sequential process that ends in the malignant transformation of cells. Tumor initiation and progression are often stochastic and dynamic and give rise to cells with diverse mutations and proliferative activity. Additionally, tumor cells are under continuous selection pressure in order to achieve a proliferative advantage and avoid elimination by control mechanisms. Together these processes promote the formation of heterogeneous tumors with genetically distinct subpopulations, called intra-tumor heterogeneity. This is in contrast to inter-tumor heterogeneity caused as a result of patient-specific factors (Burrell et al., 2013). A main driver of intra-tumor heterogeneity is genomic instability that provides the basis for genetic diversity within a tumor. Increased mutational burden due to genomic instability is caused by defects in DNA maintenance mechanisms (Negrini et al., 2010). However, genomic instability alone is not sufficient for establishing a heterogeneous tumor, since it could also lead to the outgrowth of individual clones with the highest proliferation rate. Clonal evolution is necessary to shape a tumor into a tissue with diverse clones each providing the possibility to adapt to changing environments (e.g. drug treatment or hypoxia). Due to spatial and temporal differences in the tumor environment, the distribution or the abundance of subclones within a tumor can differ. Spatial heterogeneity refers to genetically distinct subclones within the primary tumor or across metastasis. Genomic information from single cell nuclei sequencing of breast cancer biopsies revealed the presence of three distinct subpopulations, which arose from punctuated expansion of clones (Navin et al., 2011). On the other hand, temporal heterogeneity refers to changes in the clonal composition of a tumor over time induced by changes in the selection pressure. These changes are in most cases anti-tumor drug treatments, which provoke adaptation responses by the tumor. Longitudinal sampling of non-small cell lung cancer (NSCLC) over treatment and treatment free intervals, showed appearance, loss and reappearance of resistant subclones (Sequist et al., 2011), signifying the

tumors adaptation towards treatments due to the presence of drug-tolerating or resistant cells.

1.7. Drug treatments against cancer

Drug therapies against cancer aim at targeting properties of tumor cells that make them more susceptible than healthy cells. Among these properties are the disruptions in DNA repair mechanisms and chromosomal maintenance and the increased genetic instability and proliferative activity, which are exploited by chemotherapeutic drugs. Classes of chemotherapeutics are alkylating agents (bind covalently to DNA) (Damia and D'Incalci, 1998), antimetabolites (inhibition of DNA and RNA synthesis) (Parker, 2009), anti-microtubule agents (inhibition of microtubule polymerization and depolymerisation) (Rowinsky and Donehower, 1991) and topoisomerase inhibitors (prevent unwinding or replication of DNA strands) (Pommier, 2013). Since chemotherapeutic drugs only aim at a higher susceptibility of tumor cells compared to normal cells, treatments are often accompanied with strong side effects. The development of molecular target therapies provides a more specific therapeutic intervention. The dependencies of tumor cells on specific driver mutations offer great therapeutic potential, such as inhibition of specific oncogenic proteins. A prominent example is the treatment of myelogenous leukaemia with the small molecule Imatinib. As a product of gene translocation, myelogenous leukaemia cells express the fusion gene Bcr-Abl, resulting in a hyperactive Abl tyrosine kinase and thus increased cell proliferation. Imatinib specifically inhibits the Bcr-Abl fusion protein and thereby only affects tumor cells expressing it (Druker, 2008).

1.7.1. Drug resistance

Unfortunately, the high efficacies of chemotherapeutic and target therapies are often limited to short time windows due to the development of resistances. Tumor cells either adapt to the increased selection pressure applied by a drug or pre-existing resistant subclones evolve in the course of the treatment (adaptive resistance). Adaptive responses include transcriptional reprogramming of tumor cells to counteract drug damages. Tumor cells can compensate DNA damages induced by chemotherapeutic drugs by up-regulating the expression of genes

involved in the DNA repair machinery (Kirschner and Melton, 2010). A more generic resistance mechanism against chemotherapeutics is the expression of ATP-binding cassette transporters promoting drug efflux and thus resistance (Gottesman et al., 2002). Another example is the upregulated expression of drug targets, resulting in more targets to be inhibited and consequently a reduced drug efficacy, as suggested for Imatinib resistance due to Bcr-Abl overexpression (Barnes et al., 2005). Genomic instability favours the acquisition of new mutations and thus can potentially give rise to resistant cells (acquired resistance). Prominent genetic alterations to escape inhibition of kinases by small molecules are so called gatekeeper mutations that reduce the accessibility of the binding pocket. The inhibition of epidermal growth factor receptor (EGFR) by Gefitinib in NSCLCs is often followed by relapses due to a secondary EGFR gatekeeper point mutation (Kobayashi et al., 2005). Resistance to Gefitinib in NSCLCs was also found to be caused by an amplification of the *MET* gene, resulting in the HER3-dependent activation phosphatidylinositol-3-kinase (PI3K), which provides pro-survival signalling (Engelman et al., 2007). EGFR inhibition induces the activation of an alternative signalling route, a mechanism called oncogenic bypass.

Over the last decades intra-tumor heterogeneity evolved as one of the main drivers of drug resistance. Higher clonal diversity increases the chance of adaptation to drug treatments. On one hand, the presence of pre-existing subclones either being resistance or capable to adapt to drug treatment is more likely. Additionally, genetic diversity increases the probability of cells acquiring resistance. Resistance to EGFR inhibition in NSCLCs was described to either arise by the selection of pre-existing clones with a gatekeeper mutation or the acquisition of a gatekeeper mutation (Hata et al., 2016). How small numbers of pre-existing subpopulations cause tumor relapse by their adaptive selection and acquired transcriptional signatures was shown recently in the Navins Laboratory. Transcriptional and copy-number evolution profiling of single cells from triple negative breast cancer (oestrogen, progesterone and HER2 receptor negative) biopsies collected over the treatment showed the pre-existence of primed resistant cells, which eventually became resistance due to a reprogrammed gene expression. Taken together, tumor heterogeneity provides a selective advantage for tumors and thus a basis for drug resistance.

1.7.2. Drug combinations for cancer therapy

Treating tumors with drug combinations offers great potential to reduce the risk of resistance, especially for heterogeneous cancers. Targeting multiple subpopulations within the same tumor can prevent positive selection of resistant or drug tolerating clones and thereby prevents tumor relapse. Resistance to EGFR-inhibition in NSCLCs was hypothesized to be caused by pre-existing cloned with *MET* amplification and their positive selection during Gefitinib treatment (Turke et al., 2010). Therefore, co-targeting of EGFR and MET tyrosine kinase in NSCLC patients to prevent the outgrowth of resistant cells is explored in a clinical trial (NCT02335944). Furthermore, drug combinations increase the selection pressure on tumor cells and thus reduce the risk of acquired resistances. Allosteric inhibition of Bcr-Abl combined with blockage of the ATP-binding site has shown to improve treatment of Bcr-Abl positive tumors (Zhang et al., 2010). Since the two compounds have non-overlapping resistance profiles, a gatekeeper mutation against either of the drugs does not result in resistance against the other drug. Additionally, adaptation to drug treatments by compensatory responses such as activation of bypass signalling may be tackled by drug combinations targeting both pathways. Treatment of KRAS-mutant lung cancer with Trametinib, which inhibits mitogen-activated protein kinase (MAPK) downstream of KRAS, leads to resistance mediated by fibroblast growth factor receptor 1 (FGFR1). The combined inhibition of MAPK and FGFR1 was found to improve tumor cell death *in vivo* by nullifying the adaptive response (Manchado et al., 2016). These examples illustrate the power of drug combinations for the treatment of cancer by preventing relapses due to adaptive and acquired resistance mechanisms. However, finding multiple vulnerabilities of tumor cells and define treatments against those is difficult. One major reason is the existence of inter-patient heterogeneity that leads to the emergence of varying driver mutation and adaptation mechanisms (De Palma and Hanahan, 2012). Therefore, it would be beneficial to design personalized drug combination by identifying vulnerabilities for each patient, which can be achieved by high-dimensional profiling such as genome and transcriptome sequencing.

1.7.3. Gene expression profiling for precision oncology

Due to the vast heterogeneity within and between tumors from different patients, a lot of effort is put into finding the perfect drug for each patient. Instead of using

standardized therapies, precision oncology aims at identifying biomarkers that allow predictions on drug sensitivity. The spectrum of predictive biomarker reaches from single genes to entire pathway networks and their identification mainly relies on high dimensional data sets. Gene expression profiles from various disease models and patient biopsies under perturbed and untreated conditions have proven to be powerful in understanding drug responses and predicting efficacies. Detecting the expression of marker genes can be used for deciding on treatments with target drugs, such as Her2 overexpression for EGFR-inhibition in breast cancer (Nicolini et al., 2018). Gene expression data is also widely used to predict pathway activity in cancer cells by testing the enrichment of genes sharing biological functions (Subramanian et al., 2005). Determining pathway signatures for cancer cell lines was applied to predict drug sensitivity and thus therapies decisions can be guided by pathway activity inferred from gene expression data (Bild et al., 2006). By integrating large data repositories from hundreds of cell lines together with their sensitivity against hundreds of drugs, better predictions on drug sensitivity can be achieved. Correlations between drug sensitivity and genomic aberrations and gene expression were used to build models to predict treatment outcome (Iorio et al., 2016; Menden et al., 2018).

The comparative analysis of gene expression data of cells from before and after drug exposure can be applied to understand the mode of action (MoA) of drugs and infer drug specific signatures. The up- or down-regulation of pathway activities upon perturbation gives the compounds' effects on cells, called drug signature. The LINCS database (former cMap) comprises 1.3 million gene expression profiles from cell lines exposed to 42,080 different perturbants (Subramanian et al., 2017). The large number of connection between drugs and gene expression in the database finds wide application in drug repositioning by matching signatures between drugs (Iorio et al., 2010; Keiser et al., 2009). Furthermore, knowledge about drug signatures was used to predict drug treatments based on finding drugs-disease associations. The underlying hypothesis is that a disease specific gene expression signature is reverted by a drug with an opposing signature and thus the disease phenotype is reverted (Iorio et al., 2013). Gene expression data from 100 diseases was used to predict new therapeutic intervention by finding drugs with opposing signatures (Sirota et al., 2011). By defining drug resistance phenotypes and querying for drug signatures that revert the resistant state, it was possible to define

drug combinations nullifying the resistance mechanisms in cancer (Hassane et al., 2010; Wei et al., 2006). The comparative analysis of gene expression signatures from untreated and treated cells offer great potential for defining new treatment strategies against cancer. Consequently, the development of new technologies, which allow screens of high number of drugs and drug combinations with gene expression based readout, is of great interest and one of the main goals of the presented work.

2. Objectives and outline

The main objectives of this thesis were to establish microfluidic technologies that allow the production of chemical diverse combinations in picolitre-sized droplets and their multiplexed analysis using DNA based barcoding and RNA-Seq.

In order to achieve these goals, my first objective was to establish a microfluidic workflow for the generation of combinations from injected reagents controlled by a valve-based module.

To realize the multiplexed analysis of combinations in droplets, the second aim was to integrate a DNA-based barcoding approach into the pipeline and show that it is applicable for the identification of specific droplet populations.

The developed methods for the achievements of these first aims are described in **Chapter 3**

The third objective was to further increase the throughput of the valve-based generation of combinations in droplets by establishing a system that combines compounds from a valve-module and a 96-well plate to produce chemically highly complex droplet libraries.

The fourth objective of this thesis was to combine the microfluidic system with a barcoding approach to analyse the effects of drug combinations on gene expression of cancer cells in a highly multiplexed way.

Results for both objectives are described and discussed in **Chapter 4**

Finally, my aim was to apply the established workflow to perform screens of anti-tumor drug combinations and analyse their effects by gene expression. The preliminary results of two screens are described in **Chapter 5**

R E S U L T S

3. Combinatorial barcoding of droplet content for multiplexed readouts

3.1. Introduction

This chapter describes a novel microfluidic pipeline for on-demand production of picolitre-sized droplets containing sample combinations at high-throughput. Production of emulsions with high chemical complexity is of great interest, since it will enable screening of drug combinations with low sample consumption while having single cell resolution. A bottleneck for generating chemically diverse droplets, however, is the need for changing the aqueous phase injected into a microfluidic device for droplet production. The identification of droplet contents from a chemically complex emulsion was so far achieved by fluorescence-based barcoding or spatial order of droplets resulting in limited scalability and throughput. Various approaches for the generation of chemically diverse droplets have been presented, which have been limited to the encapsulation of individual drugs (Brouzes et al., 2009; Clausell-Tormos et al., 2010; Gielen et al., 2015; Miller et al., 2012). In previously described valve-based system, combinations were generated in large droplets (~0.5 µl), which were stored sequentially in a tubing to maintain spatial order (Eduati et al., 2018). In this chapter, the first part of the study is described, where the microfluidic workflow as presented by Eduati *et* al. was adapted and integrated into a new pipeline in order to generate combinations in an emulsion of picolitre-sized droplets. For the identification of droplet content, we introduced a DNA-barcoding approach to encode each combination with a set of two barcodes allowing multiplexing of several distinct conditions. We validated the combinations generated with our pipeline using fluorescence measurements as well as NGS based read-outs. Possible applications of this system are combinatorial drug screenings in which droplets containing potent combinations are enriched by

sorting (e.g. based on an apoptosis assay of co-encapsulated target cells) and hits are identified by sequencing of the corresponding barcodes. We performed a mock screen in which a specific droplet population was enriched based on fluorescence signals and hits were identified based on their barcodes. Additionally, we simulated a more realistic screening result, in which barcode containing droplets were introduced at different amounts and assessed the sensitivity of the system to differentiate small changes in droplet abundance.

3.2. Microfluidic pipeline for generating sample combinations in droplets

In order to achieve a continuous and automated production of droplets containing distinct combinations of two components, a microfluidic pipeline consisting of three main modules was established **(Fig. 3.1**): (1) A valve-module with 16 inlets directing continuously injected compounds either to the waste or the outlet of the microfluidic chip **(Fig. 3.1A-1, 3.1B)**. The computer-controlled Braille valves can be opened and closed in defined sequences and time intervals, allowing rapid switching (approx. 200 ms) between different injected liquids. The simultaneous opening of two valves resulted in a combination of the two components injected via the opened valves. Thus, 49 combinations (7x7; samples A-G combined with samples 1-7) were generated. Between individual combinations, two valves for oil injection were opened, generating an oil spacer that separated two defined plugs. These plugs were then transferred via tubing to the second module. (2) A drop-maker device where plugs with combinations from module 1 can be combined with cells or reagents and encapsulated at a flow focusing junction into picolitre-sized droplets (**Fig. 3.1A-2, 3.1C**). This workflow enabled the generation of 49 combinations and their encapsulation into droplets together with single cells or further reagents. (3) A third module was integrated into the pipeline, in order to prevent cross-contamination from remaining liquids in channels of the valve-module and drop-maker. The two outlets of the drop-maker device were connected with the two valves of module 3 such that the produced droplets could either be sent to a collection or waste tube (**Fig. 3.1A-3, 3.1D**). The sorting decision was depending on the fluorescence signals of incoming plugs, monitored on the drop-maker, and collection parameters (fluorescence threshold and time windows).

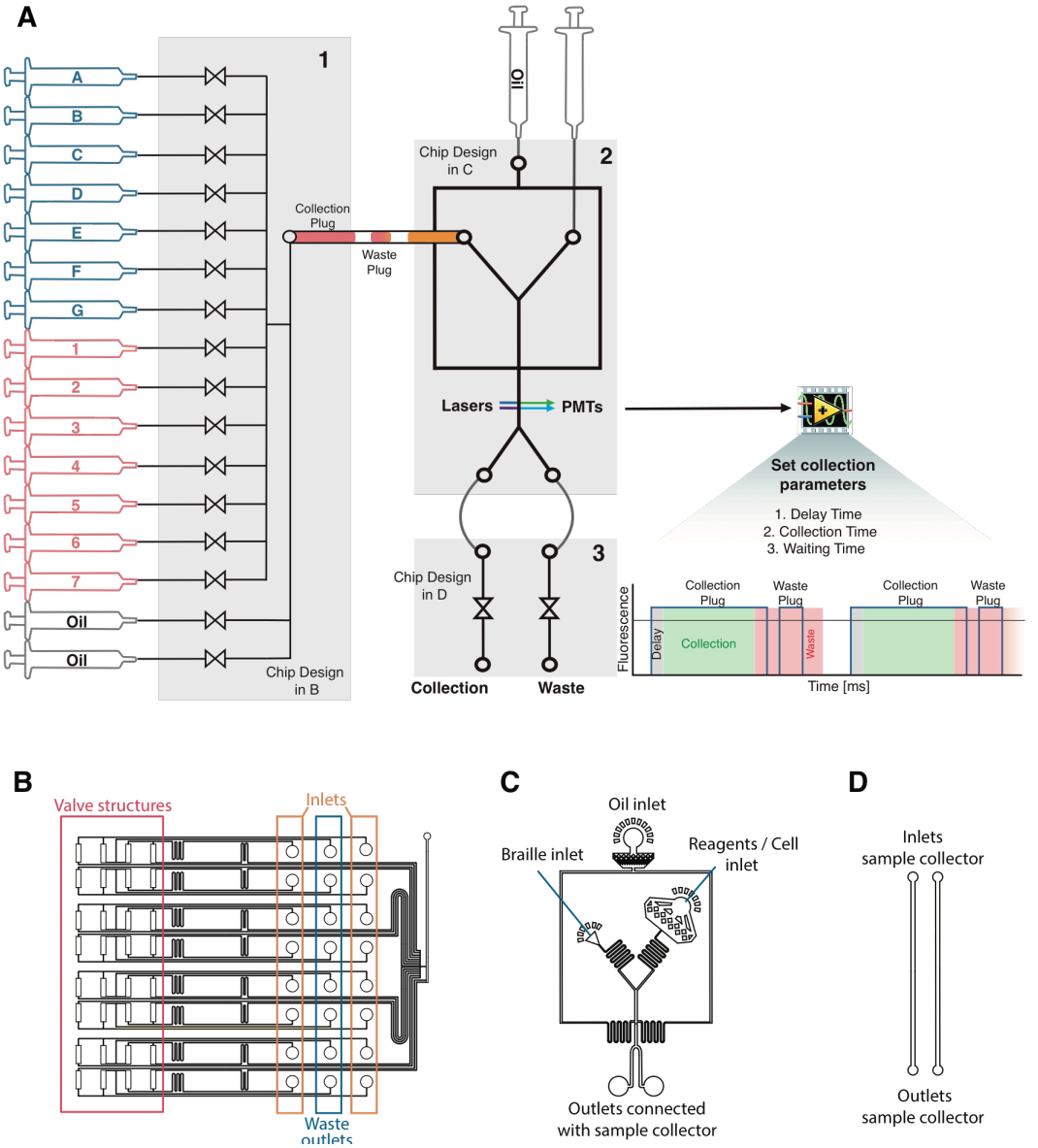

Fig.3.1 Microfluidic pipeline for generating combinations in droplets

(A) The valve-based system for generating droplets containing a mix of two compounds comprised three modules (microfluidic devices). **(1)** A valve-module with 16 valves was made up of a braille display and a microfluidic chip that is aligned with its valve structure on top of the braille pins. Opening two valves simultaneously generated combinations of two injected solutions (A to G combined with 1 to 7). **(2)** Combinations were transferred in a small piece of tubing to the drop-maker module and encapsulated into droplets along with cells or reagents. **(3)** A sample collection module consisting of two valves allowing to direct droplets to the waste or collection tube. The collection and waste valve of module 3 was operated by a LabVIEW software. Fluorescence signals of injected plugs containing combinations were processed by the software to initiate a time series (1. Delay Time, 2. Collection Time, 3. Waiting time) of different steps, required to collect droplets without cross-contamination. When signals from a sample plug were detected, the sequence of delay, collection and waiting time was started ensuring collection of droplets from sample plugs (green box) whereas droplets from washing plugs are discarded (red box). **(B)** Layout of the microfluidic chip used for the valve-module. Rectangular channels (red box) are aligned on pins of a braille display to direct solution injected into the inlets (orange boxes) to the waste outlets (blue box) or chip outlet. **(C)** Design of the drop-maker chip used for module 2. **(D)** Two valving channels of sample collector chip allowing to direct drops to the waste or collection tube.

Measured fluorescence signals of incoming plugs were processed by an in-house LabVIEW software that controlled collection and waste valves. When the fluorescence signal reached a pre-defined threshold a series of delay-, collection- and waiting-times were activated. During the delay time (~100-500 ms, grey box), droplets were sent to the waste although the set fluorescence threshold was reached, allowing the droplet generation to equilibrate. Following the delay, a collection time was set to collect droplets from the collection plug. As a result, the collection valve was open for a pre-defined collection time (2s to 8s, green box). During the waiting time the system sent all produced droplets to the waste irrespective of the fluorescence signal (5-8 sec, red box). The waiting time window was used to remove remaining liquids of previous samples by the passage of the washing plug generated from the subsequent sample. Once the waiting time was over, the system was ready for the collection of the next plug by repeating the cycle.

3.2.1. Design of the valve-module and sample collector

The valve-module and sample collector described in 3.1 consisted of a braille display and a microfluidic chip. For the valve-module the chip in **Fig. 3.1B** was used. Briefly, a PDMS chip with rounded channels was bonded to an elastic PDMS membrane. The wide rectangular structures (valve structures) were aligned on top of the pins of a Braille display. Reagents connected to the inlet ports were continuously injected and reagents were sent to the waste outlets in the default mode. In order to direct injected reagents to the drop maker, the waste channel was blocked by moving the corresponding braille pin upwards, while the pin blocking the flow towards the drop maker was moved down. Thus, one valve can be defined by a set of two pins, which direct one injected reagent either to the waste outlet or the chips outlet connected to the drop-maker. This process was automated by a LabVIEW software that allowed to run defined series of sequential valve opening times. The working principle of the sample collector was the same except that four pins were aligned below one of the channels of the chip design shown in **Fig. 3.1D**. Opening and closing of channels was used to direct droplets to the collection or waste tube. More details can be found in the Material & Methods chapter of this thesis.

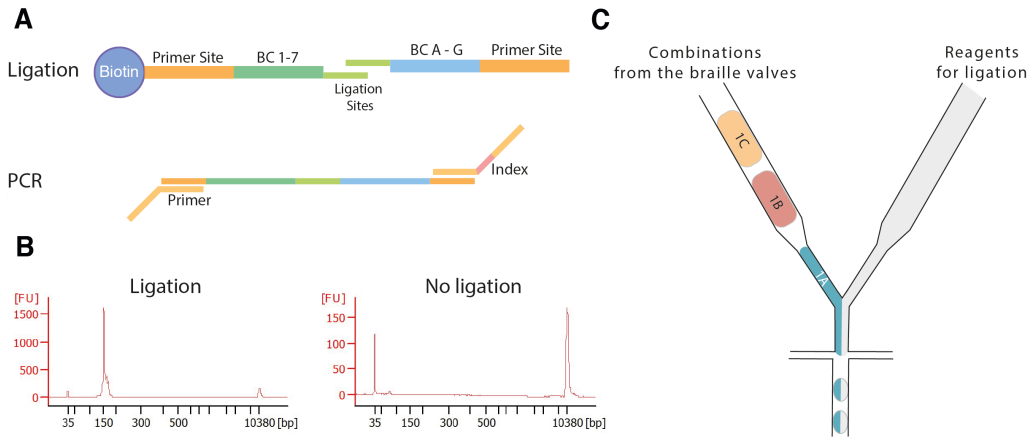

Fig.3.2 DNA-based barcoding of droplet content

(A) Biotinylated barcodes are hybridized over a ligation site with a second set of barcodes and subsequently ligated. This results in a barcode combination with two primer sites that can be amplified by PCR. Fragments are inked to sequencing adapters during PCR resulting in Amplicon-seq libraries in a one-step protocol. **(B)** DNA high sensitivity Bioanalyzer results of ligated and non-ligation barcodes after PCR. Ligation and amplification results in 150 bp fragments. **(C)** Generation of droplets at the flow focussing junction of the drop-maker containing a combination of two barcodes (1A, 1B, 1C, …, 7E, 7F, 7G) and reagents for ligation. Barcode combinations are injected from the valve-module in form of distinct sample plugs spaced out by oil. The volume of sample plugs and thus the number of droplets produced for each combination depends on the valve opening times.

3.3. DNA-based barcoding of combinations in droplets

A challenge in handling chemically diverse droplet populations is to maintain the information about droplet content. A possibility to overcome this limitation is to uniquely label each droplet according to its content. This was so far achieved by using fluorescent dyes, which have limited multiplexing capabilities. In order to encode and read out information of droplet content, we have established a barcoding approach in which two DNA barcodes (BCs) form one functional barcode in a ligation reaction. Biotinylated double stranded oligonucleotides with a forward primer site and a 10 bp barcode (BC-1 to BC-7) were joined over single stranded ligation sites with oligonucleotides comprised of a 10 bp barcode (BC-A to BC-G) and a reverse primer site **(Fig. 3.2A)**. After droplet breakage, barcodes were purified and amplified using primers with Illumina sequencing adapters yielding in amplicon sequencing libraries. Upon ligation, PCR yielded a strong amplification of the 150 bp long fragments, whereas non-ligated barcodes did not result in any detectable amplification **(Fig. 3.2B)**. The barcoding approach was integrated into the microfluidic pipeline by injecting BC-1 to BC-7 and BC-A to BC-G into the valve-

module and combining the barcode species with each other. Plugs containing barcode combinations were subdivided into picolitre-sized droplets, together with reagents for ligation, using a flow-focusing junction on the drop-maker chip **(Fig. 3.2C)**.

3.4. Pipeline generates droplets with uniform size and contents

For demonstrating the working principle and robustness of the microfluidic pipeline, a valve-module with 16 inlets was used and generated droplets were analysed by fluorescence signals. Cascade Blue (A-G) injected over seven valves was combined with injected fluorescein (1-7) in order to measure how the modules combine the two compounds. An important parameter for droplet generation in a microfluidic device is the monodispersity since a uniform size reflects a stable droplet production with uniform content. To assess the monodispersity of droplets generated from sample plugs, we measured the width of the fluorescence signals from droplets of 49 combinations generated by the valve-module (**Fig. 3.3A**), and no major fluctuation in droplet size was observed. This observation was quantified by calculating the coefficients of variation for each group of seven combinations, which yielded values ranging from 5.3% to 6.5% **(Fig. 3.3B)**. Measuring the fluorescence signal of droplets allowed us to determine whether both compounds from the valve-module were mixed at equal ratios and if they were encapsulated homogeneously with the other reagents injected into the drop maker. Blue fluorescence intensities of droplet generated from always combining green dye (inlets 1-7) with blue dye (inlets A-G) showed variations ranging from 6.6% to 12.7% for groups of seven combinations **(Fig. 3.3C)**. The observed variation of the median fluorescence intensities between all seven groups was 3.3%. In order to further evaluate the valve-based mixing of two components, syringes were supplemented with DNA barcodes. All barcode combinations of 1-7 with A-G were generated, collected separately and ligated. The quantification of barcode combinations by qPCR showed a low level of variation (CVs between 2.2% and 5.1%) except combinations of BC-1 with BC-A to BC-G (CV of 7.7%) most likely due to pipetting errors. These results imply a stable and uniform encapsulation of all components into picolitre-sized droplets **(Fig. 3.3D)**.

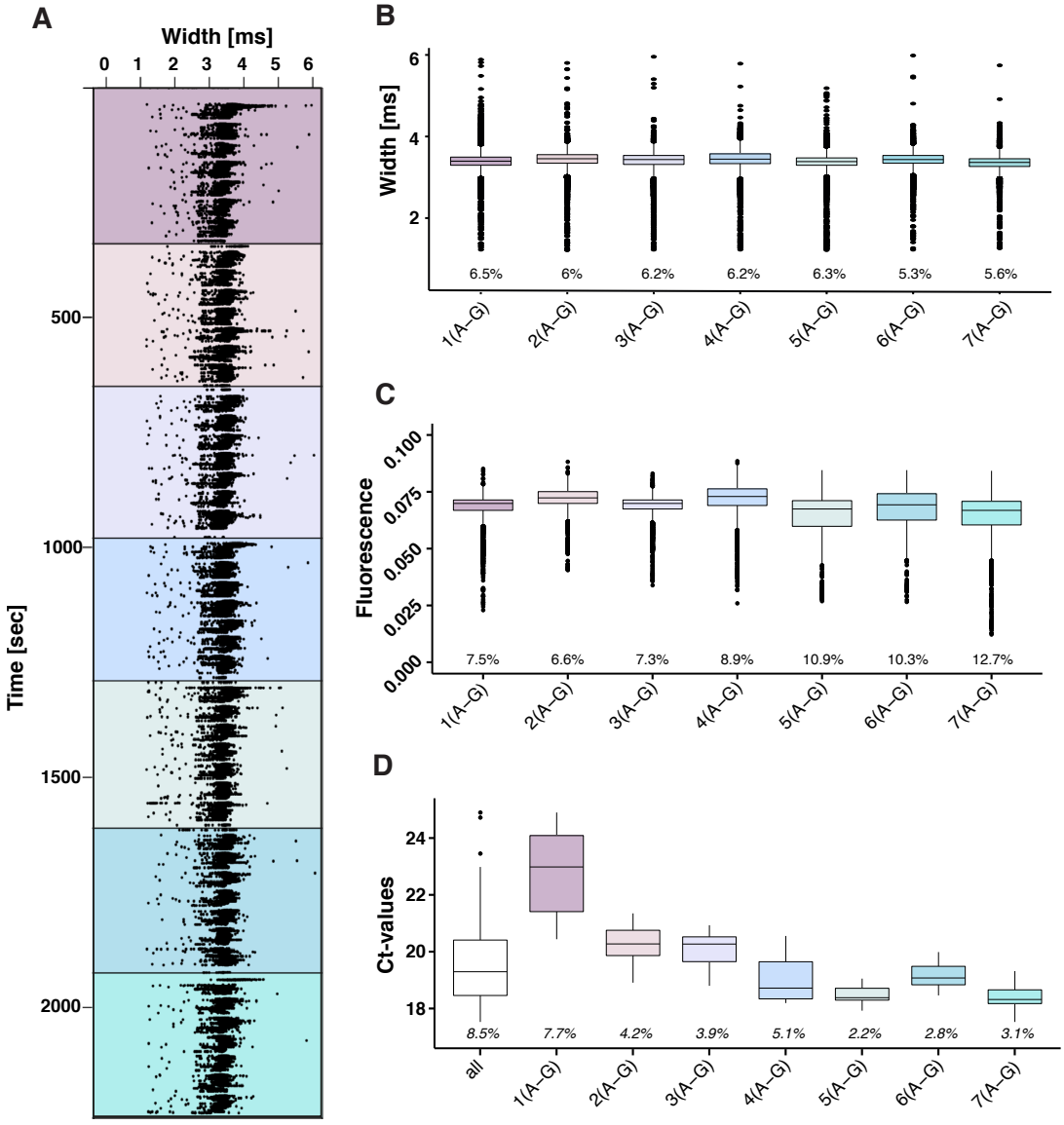

Fig.3.3 Generation of combinations in droplets using braille valves

(A) Measured fluorescence signal width in milliseconds [ms] over 49 combinations with each dot representing a droplet and the coloured boxes indicating a set of 7 combinations. **(B)** Boxplots of signal width grouped by seven combinations according to the colour scheme in A. *Numbers* show coefficient of variations (CV) of signal width. **(C)** Blue peak fluorescence intensities of droplets generated from combinations (1-7 + A-G). Combinations were grouped according to the colour scheme in A and *CVs* were calculated for each group. **(D)** Ct-values of 49 barcode combinations produced from the valve-module. Barcodes were collected separately and then amplified. Ct-values were grouped into seven combinations to calculate *CV values.*

3.5. Sample collector and washing steps

The microfluidic pipeline included a sampler collector module, which was integrated to collect specific subsets of droplets. It was necessary to perform washing steps between each combination to prevent cross-contamination. After a combination was generated, valves to produce the next combination were opened for 3 sec resulting in a waste plug (as it could contain remains from the previous combination), which was followed by an oil spacer. Once the waste plug flushed the channels, the same valves were opened to produce the actual collection plug, without any contamination from the previous sample composition. This resulted in alternating sample and waste plugs spaced out by oil. The sample collector was, thus, set to only collect droplets produced from sample plugs by applying time windows for opening collection or waste valves. To test the collection and disposal of distinct droplet populations, sample plugs with fluorescein (Green channel) and Cascade blue (UV channel) followed by waste plugs additionally labelled with Alexa 594 (Orange channel) were generated using the valve-module **(Fig. 3.4A)**. We aimed to collect only the droplets generated from green sample plugs and discard droplets generated from orange plugs by setting time windows for delay, collection and waste (coloured boxes in Fig. 3.4A). Droplets from the collection and waste tube were re-injected into a microfluidic device to measure fluorescence intensities of the two fractions. Droplets from the collection tube were negative in the orange channel, whereas droplets from the waste were positive, signifying a robust and efficient collection of desired droplets generated from sample plugs **(Fig. 3.4B)**. The waste contained additionally orange negative droplets, due to the disposal of droplets from collection plugs during delay times. To assess the purity of sample plugs, we generated combinations from seven inlets containing Cascade blue with six inlets containing water and one inlet containing fluorescein. Thus, by combining one Cascade blue inlet, six blue washing and sample plugs and one double positive (green and blue) washing and sample plug was produced **(Fig. 3.4D**). The summarized fluorescence intensities of all plugs showed strong green signals for the plugs expected to be positive, whereas the subsequent washing plugs and the plugs expected to be negative showed only slightly increased or background signals, respectively **(Fig. 3.4C)**. This demonstrates that a washing step of 3 sec was sufficient to remove all residual liquids that remain from the previous sample composition in the channels.

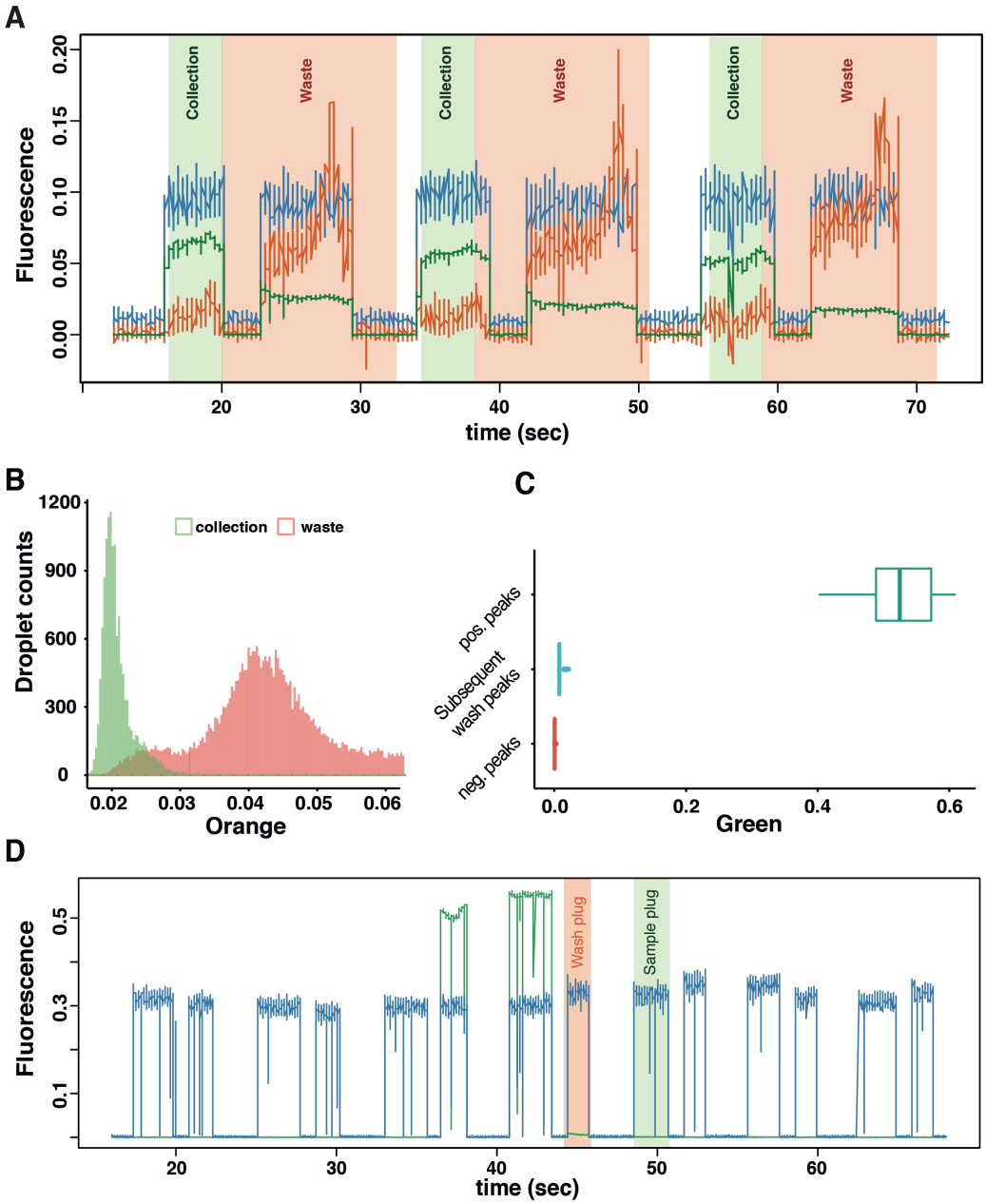

Fig. 3.4: Validation of the sample collector

(**A**) Fluorescence peaks of green positive (Fluorescein + Cascade Blue) and orange green double positive (Alexa594 + Fluorescein + Cascade Blue) plugs. Collection parameters were set to collect droplets generated from green plugs (green box) and discard droplets from orange plugs (orange box). (**B**) Histogram of droplet intensities in the orange channel. Green and red counts correspond to measured intensities of droplets from the collection tube and waste tube, respectively. (**C**) Quantification of green fluorescence intensities of pos. green peaks, the subsequent washing peaks and the remaining peaks (neg. peaks). **(D)** Representative fluorescence peaks of sample (green box) and washing (orange box) plugs and green peaks every fourth plug. Shown peaks correspond to a cycle of seven combinations.

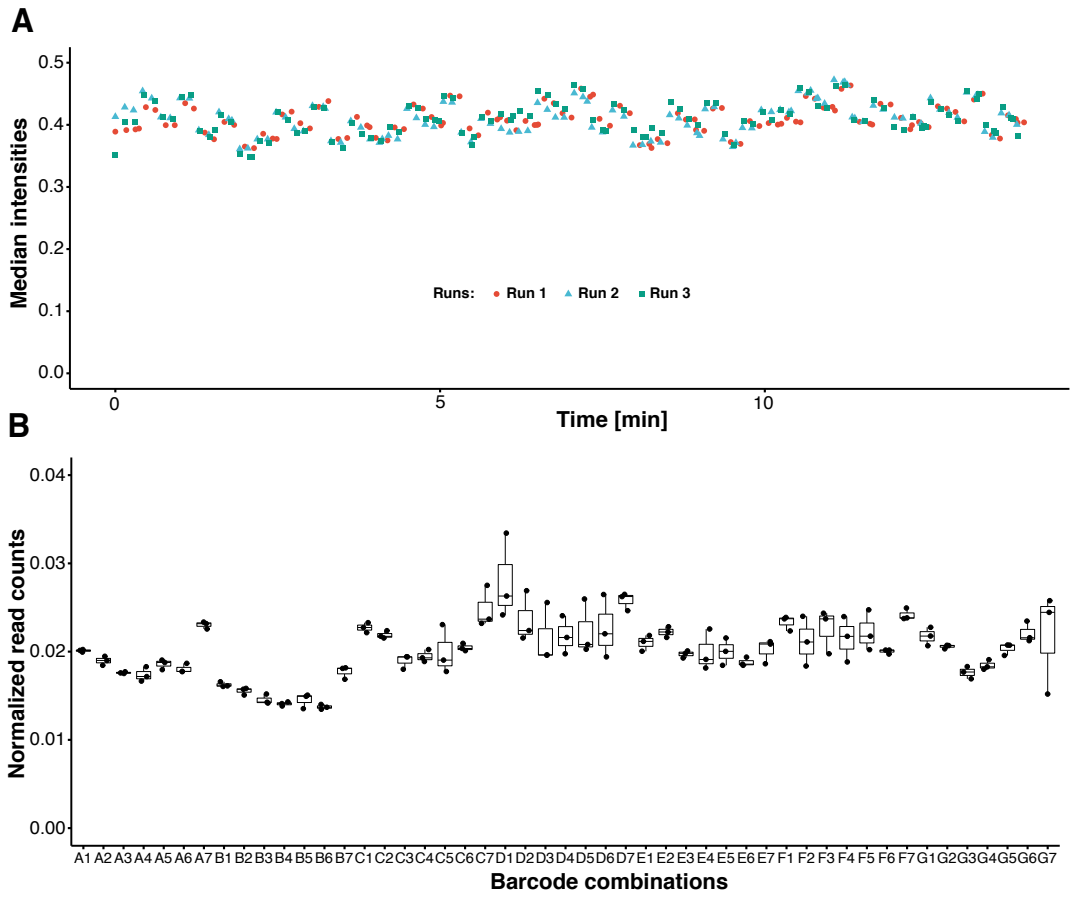

Fig. 3.5: Barcode combinations generated in droplets

(**A**) Median fluorescence intensities of sample plugs from 49 barcode combinations over 3 replicates. (**B**) Normalized read counts of 49 barcode combinations for 3 replicates. Read counts were normalized by the total library size per replicate.

3.6. Generation of barcode combinations in droplets

For generation of droplets containing distinct combinations, we used a set of 7x7 DNA-barcodes, resulting in 49 combinations of barcodes which were read-out using NGS. Syringes filled with BC-A to BC-G were labelled with Cascade Blue for measuring plugs and triggering the collection of droplets. Since only one set of barcodes was labelled with a blue dye, measuring fluorescence signals was also used as a quality control for a mixing at constant ratios between labelled and non-labelled barcodes. Valves to generate one barcode combination were opened for 5 sec and droplets containing barcodes and ligation reagents were collected for 4 sec, resulting in a total production time of 12.25 min for 49 combinations with ~2100 droplets per combination (for droplets of 800 pl in size). The fluorescence data of barcode combinations showed that the median intensities of plugs for all

combinations were stable between the three replicates. The median intensities over all the plugs within each of the three technical replicates were 6.5%, 6.6% and 6.5%, indicating that the ratios at which two barcodes were mixed were stable over the experiments. Barcode combinations were sequenced and reads were demultiplexed based on the input sequences from the barcode library. Read counts for each barcode were normalized by the total number of reads per replicate. For the majority of barcode combinations only low level of variation was observed between replicates (Median CV=3.9%), conforming reproducibility between replicates **(Fig. 3.5B)**. However, the abundances of different barcode combinations showed an almost 2-fold difference in the read-count medians. This can be circumvented by comparing barcode abundances to a reference like the barcode numbers before sorting-based enrichment (see chapter 3.7).

3.6.1. Collection of DNA barcodes of varying collection times to mimic differences in barcode abundance

In order to assess the sensitivity of detecting barcodes at different abundances, we performed experiments with varying collection times for different barcodes. For a set of seven barcode combinations the collection times were lowered manually from 8 sec to 2 sec by increments of 1 sec in the sample collector software **(Fig. 3.6B)**. Apart from fluorescence intensities of sample plugs, the sample collector software recorded the time windows of droplet collection **(Fig. 3.6A)**. The number of data points in each collection time window reflects the number of droplets collected for each combination. Since droplets were produced at 520 Hz and data points for sample plugs were acquired at 100 Hz (upstream of the droplet production), one data point corresponds to 5.2 collected droplets. Hence the number of data points was used to verify, if correct numbers of droplet for all 49 barcode combinations over 3 replicates were collected **(Fig. 3.6B)**. Only in one case (Barcode combination D7) we observed a strong deviation in the number of collected droplets, most likely being caused by a false activation of the collection time. The reduced droplet collection was also reflected by a reduced read-count in the sequencing data and hence this data point was excluded from the analysis.

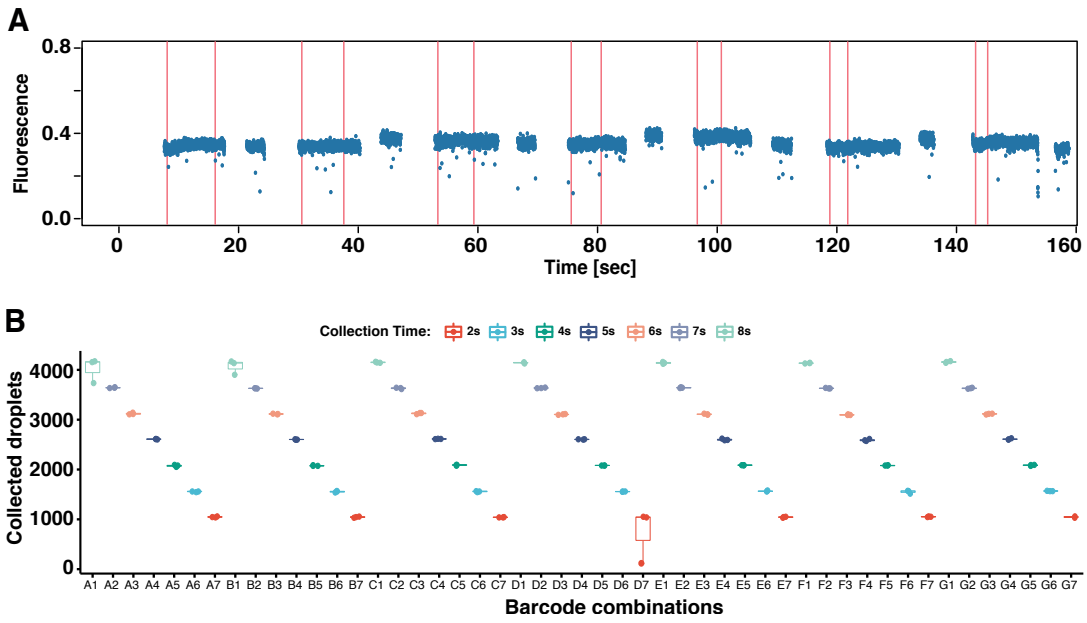

Fig. 3.6: Decreasing collection times for pairs of seven barcode combinations

(A) Fluorescence signals of plugs containing barcode combinations and collection time windows (red lines) applied by the sample collector software. (**B**) Number of droplets collected during each time window for all 49 barcode combinations (n=3). The number of droplets for each combination was determined from the number of recorded data points for each combination and the frequency of droplet.

3.6.2. Read-counts in barcode libraries recapitulate number of collected droplets

Barcode libraries of varying abundances were sequenced to determine if different amounts of barcoded droplets could be quantified using DNA-based barcoding and NGS. From each replicate two samples were prepared and sequenced. Demultiplexed read-counts for each replicate were normalized by total read-counts. To assess whether the number of generated droplets correlated with read-counts, a set of seven barcode combinations (A1-A7, B1-B7 etc.) was normalized by the read-counts for barcodes collected for 5 sec. Fold changes of barcode abundances compared to a collection of 5 sec was in good agreement with the collection times for most of the barcode combinations **(Fig. 3.7A)**. Only in three cases a reduction in the collection time did not result in a reduced fold change (D1 and D2, E3 and E4 and C6 and C7). The good agreement between collected droplets and read-counts was confirmed by the fold changes for each collection time summarized across all

barcodes, which consistently showed an increased read-count for increasing amounts of collected droplets, apart for increasing the collection time from 5 to 6 sec. **(Fig. 3.7B)**. Nonetheless, these results signify a good sensitivity since differences of approximately 520 droplets were detected in most cases with good confidence.

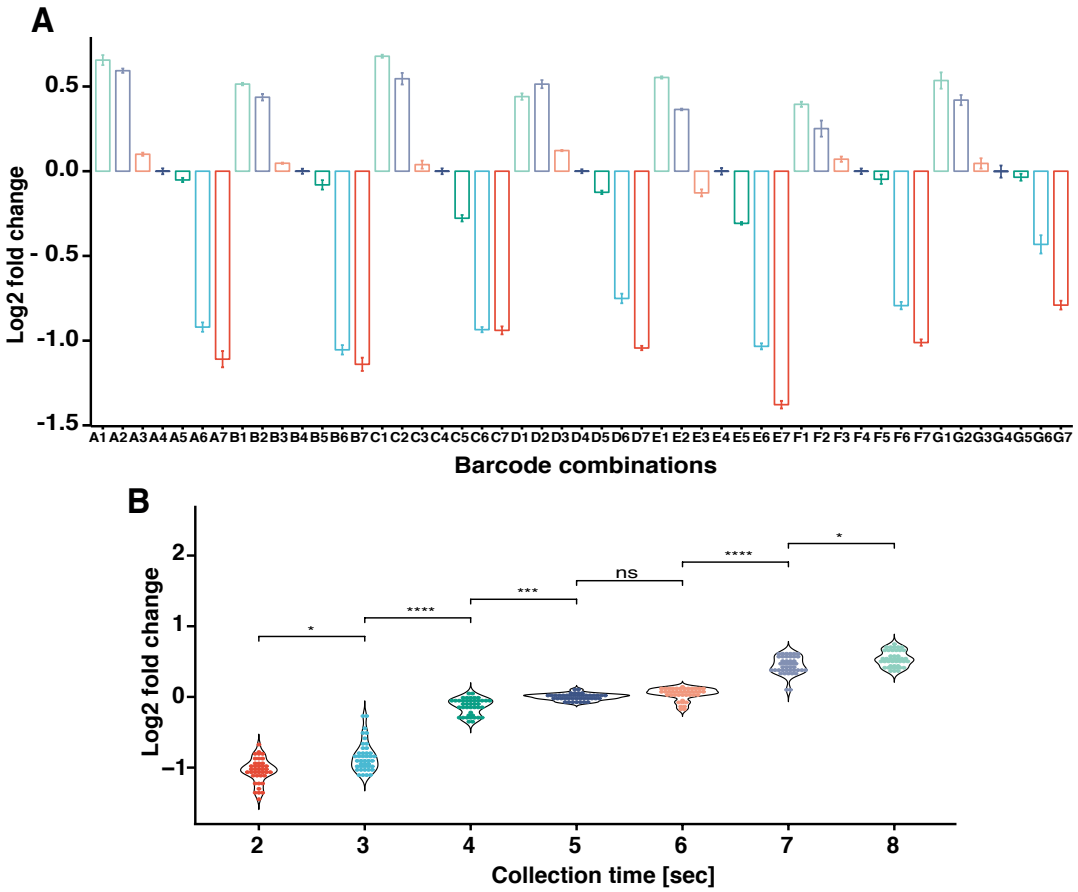

Fig. 3.7: Barcode libraries collected at varying collection time

(**A**) Log fold change of read counts for varying collection times over 6 technical replicates from 3 different samples. Barcoded droplets were collected for different time periods decreasing from barcode 1 (8s) to barcode 7 (2s). Read counts were normalized by the mean read count at 5 sec for each set of 7 combinations (**B**) Summary of the fold changes in read counts for different collection times, compared to 5s collection. Statistical significance for every increase in the collection time was determined using pairwise Wilcox test and p-values were adjusted according to Benjamini & Hochberg (ns: not significant, *: $p<0.05$, **: $p\leq0.01$, ***: $p\leq0.001$, ****: $p\leq0.0001$).

3.7. Enrichment of specific droplets encoded by barcode combinations

Fluorescence activated droplet sorting (FADS) can be applied to enrich specific droplet populations based on their fluorescence signals and thereby select the ones with specific phenotypes (e.g. droplets containing apoptotic cells). Sorting experiments with chemically diverse droplets encoded by DNA-barcodes would allow highly multiplexed enrichment of all rather than only one sample combination (e.g. drug combination) fulfilling the selection criteria. For testing the ability to enrich specific phenotypes and to determine the corresponding droplets composition by sequencing, 49 barcode combinations were generated in droplets. Apart from the syringes containing BC-A to BC-G labelled with Cascade Blue, we labelled the syringe containing BC-3 with fluorescein. As a result, 7 out of the 49 combinations (3A to 3G) were stained with a green dye, mimicking a particular phenotype **(Fig. 3.8A)**. The generated droplets from 49 combinations were injected into a droplet-sorting device and sorted based on their fluorescence signal **(Fig. 3.8B)**. Since we aimed for enriching green positive droplets, we applied a sorting gate around droplets with high green fluorescence signal **(Fig. 3.8C)**. In order to determine the level of enrichment in the sorted population, we sequenced barcode libraries prepared from droplets before sorting and after sorting and compared the read counts. The fold change in barcodes between the unsorted and sorted samples showed an increase for barcode combinations 3A to 3G (FDR<0.05) whereas all other barcode combinations were depleted **(Fig. 3.8D)**. On average, barcode combinations encoding green positive droplets were significantly enriched compared to green negative droplets ($\Delta log2 = 5.3$; $p < 2.2e^{-16}$, Student's t-test). This demonstrates that the described workflow can be applied to specifically enrich all sample combinations fulfilling particular phenotypic selection criteria (i.e. hits) by sorting based on fluorescence and these hits can subsequently be identified based on DNA-barcodes.

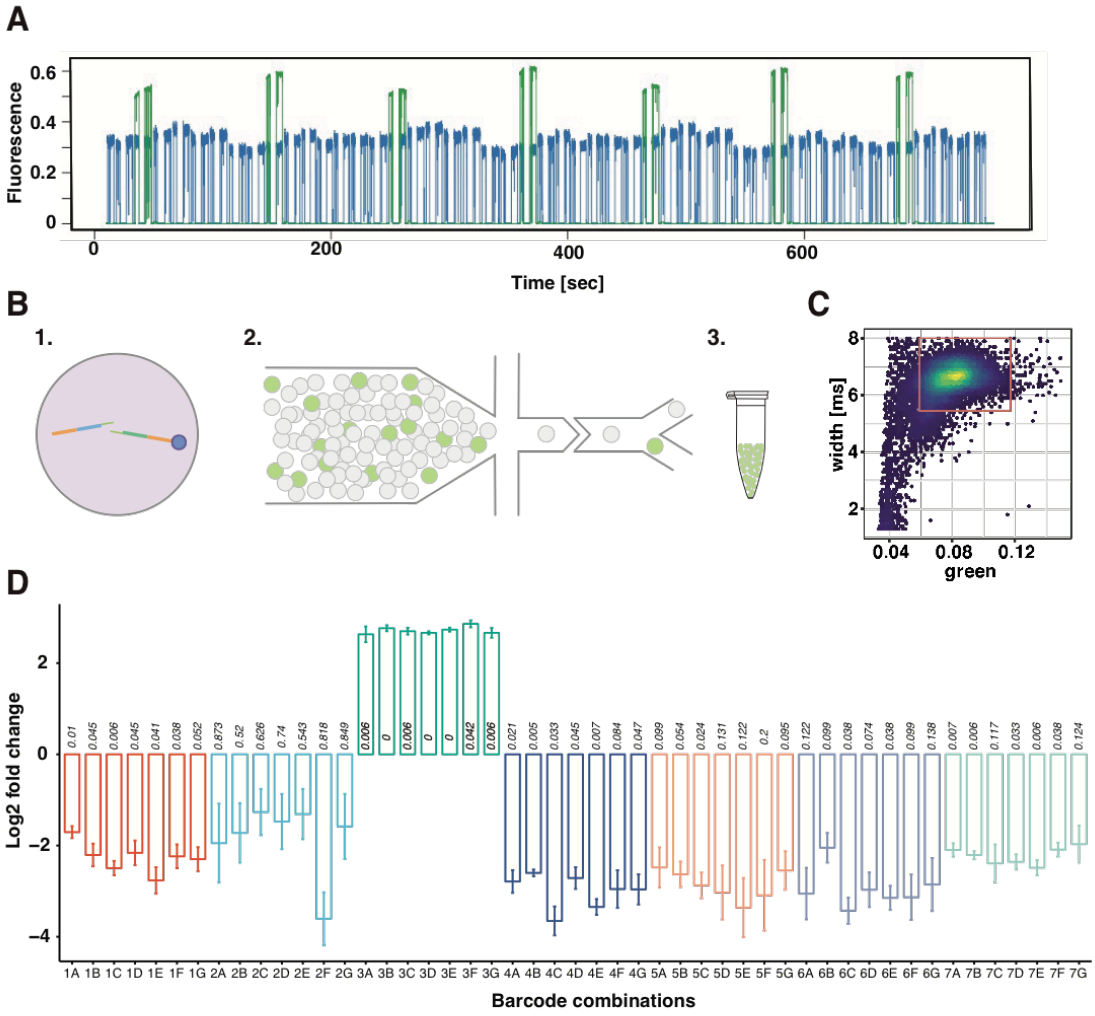

Fig. 3.8: Enrichment and demultiplexing of specific droplet populations

(A) Sequence of fluorescence signals for 49 barcode combinations. Plugs containing BC-3 were additionally labelled with a green dye resulting in green positive peaks for every 7th combination. **(B)** Microfluidic setup for the enrichment of specific droplet populations encoded by barcode combinations. **(1)** Each barcode combination was encapsulated together with ligation reagents. **(2)** Droplets were re-injected into a sorting device and fluorescence-based sorting was performed to enrich droplets positive in the green channel. **(3)** Green positive droplets were collected and used to prepare barcode libraries for sequencing. **(C)** Fluorescence data acquired during sorting of green positive droplets. Red square indicates the gating applied for sorting green positive droplets. **(D)** Log2 fold-changes in barcode abundances comparing read-counts from before and after sorting. Read counts were normalized by the corresponding means from the unsorted samples. Adjusted p-values (False discovery rate) are shown next every bar in italic.

3.8. Discussion on DNA-barcoding based multiplexing of droplet content

The above described microfluidic workflow enables on-demand generation of combinations in droplets using a valve-based system. This was achieved by integrating previously described Braille valves into a newly developed microfluidic pipeline including a drop-maker and a sample collector. The pipeline facilitates the production of droplets from sample plugs containing combinations of two samples, each encoded by a distinct DNA-barcode combination. Additionally, the sample collector enables the selective collection of droplets and thereby allows washing between consecutive combinations ensuring purity of produced droplets. The modularity of the described system has the advantage that individual components can be easily exchanged for different purposes (e.g. more inlets on the drop-maker). Furthermore, malfunctions in one of the modules can be fixed by only replacing the affected device, resulting in a workflow that is easy to maintain.

3.8.1. Validation based on fluorescence signals and qPCR

By using two fluorescent dyes, the ability of the valve-module and the drop-maker to combine two injected compounds and to encapsulate those into droplets along with other reagents was tested. The signal width of the generated droplets was used to determine the level of monodispersity. According to literature, an emulsion is generally considered monodisperse if the size of droplet diameter varies by a CV less than 5% (Roberts et al., 2012), although, monodispersity has also been attributed to emulsions with CVs in droplet width of 8% (Abate et al., 2011). The measured size distribution of droplets generated from sample plugs had a CV of 6%, thus exceeding the general criteria of monodispersity and therefore droplets can be considered quasi-monodisperse. The increased variation in droplet size was caused by high numbers of outliers (see Fig. 3.3A), which mainly fall below the 1st quartile (3% of all droplets compared to 0.9% above the 4th quartile). This is likely caused by the need for spacing out individual combinations with oil, resulting in short intervals of droplet production from aqueous phases injected into the drop-maker. Therefore, the aqueous flow needs to re-equilibrate for each combination causing short periods of flow instability. Additionally, the variation measured for the fluorescence intensities of the droplets affects the apparent signal width (higher

peaks cause a greater signal width) and thereby contributed to the measured quasi-monodisperse size distribution (Clausell-Tormos et al., 2008). As a consequence the true monodispersity of the generated droplets might be better than the measured size distribution. The dilution of a fluorescence dye was used as a measure to determine the mixing of reagents by the valve-module. In some cases we observed strong variations in the fluorescence intensities of droplets from 49 combinations (CV > 10%, Fig. 3.3C). However, the median intensities of each group of seven combinations were stable over all seven groups (CV of 3.3%). This suggests that the mixing of two components in the valve-module was comparable for all combinations, but the droplet production was not stable over time. This again is likely to be caused by unstable flow regimes at the beginning of each collection plug and could potentially be overcome by increasing the droplet collection times for each combination. Furthermore, fluctuation in the trajectory of droplets within the channel can affect their position relative to the focused laser beam used for the readout, potentially causing signal variation that does not correspond to varying contents or sizes. Since droplets were produced from a laminar flow and no mixing of droplet content was performed before measuring, this can additionally affect the fluorescence intensities. Taken together, the true variation of droplet composition and size might be smaller than indicated by the fluorescence signals.

The accuracy of mixing two components was therefore also measured by qPCRs of two DNA-fragments combined into droplets. The Ct-values of barcode combinations were comparable for sets of seven combinations, demonstrating a stable and robust mixing of two components by the described pipeline. Since no internal reference for the normalization of Ct-values was present, the high level of variation observed for Ct-values of combinations from valve 1 with valves A to G is likely to reflect pipetting errors.

For the pipeline described in this chapter we used a sample collector ensuring the collection of only those droplets containing sample combinations without significant cross-contaminations. Washing the microfluidic devices by opening valves for the next combination and flushing residual liquids of the previous combinations away, primed the channels for collection of an uncontaminated subsequent combination. Discarding droplets created during the washing step was possible with a simple valve-based unit that directed these droplets to the waste. The implementation of a computer guided droplet collection unit allows

fluorescence- and time-based collection of droplets into different tubes. The established droplet collection module is a useful tool for droplet production were cross-contaminations between chemically different samples is an issue (Clausell-Tormos et al., 2010). The sample collector module can be further envisioned as an application such as automated droplet collection to selectively distribute droplets into different wells (Utharala et al., 2018) for more complex downstream analyses.

3.8.2. Multiplexing of droplets using deterministic barcoding

In order to show that the presented microfluidic workflow can generate truly distinct combinations, we generated 49 DNA-barcode combinations in droplets for sequencing. Using constant collection times for each combination, we expected to obtain even read counts for each barcode combination. However, we observed an almost 2-fold difference between the lowest and highest median read count (Range: 0.0137 to 0.0263). The observed variation in read-counts made it crucial to normalize values and compare relative barcode abundances. Since we envision performing combinatorial drug screenings with the presented technology, we anticipate comparisons between barcode abundances before and after a droplet sorting based enrichment step. Thereby we can neglect differences in barcode abundance before sorting, as we will only consider the enrichment of barcodes between unsorted and sorted samples. Additionally, the moderate variation between replicates will ensure reproducibility over several experiments enabling a more robust detection of barcode abundances. The observed variation in barcode abundance was potentially caused by amplification biases during library preparation. Consequently, the integration of unique molecular identifiers (UMIs) into the barcoding system could help to improve the readout of barcode abundance since absolute numbers of barcodes would be counted (Islam et al., 2014).

In a setting where droplets containing diverse drug combinations are enriched based on their potency to induce apoptosis, we would imagine an increase in number of droplets with different drug combinations that correlate with their apoptotic potency. To mimic such a case and to test the sensitivity of our setup to detect subtle differences in droplets numbers by sequencing barcode combinations, we collected droplets containing barcode combinations with varying time windows. A change in collection time for droplets with a specific barcode was reflected in most cases by a corresponding change in the read-counts. An increase in collection

time from 7s to 8s corresponded to a 1.14 fold increase in number of collected droplets, that was overall detected with statistical confidence, implying a good sensitivity in discriminating the potencies of potential "hits". Only in three cases the read-count data did not correspond to the number of collected droplets. Since we did not observe a deviation in the number of collected droplets for these cases (Fig. 3.6B), we assume that amplification biases caused the observed deviation in read-counts.

Furthermore, by sorting specific droplet phenotypes, we illustrated the applicability of the presented pipeline for combinatorial screenings of compounds with fluorescence-based detection of hits and their sequencing-based demultiplexing. We showed that 49 potential conditions could be screened simultaneously in a single droplet-based sorting experiment by sorting droplets of interest and identifying their contents by sequencing. The high specificity of the presented approach was reflected in significant increase of the desired barcode combinations (FDR<0.01) and strong depletion of all other barcodes.

4. Development of a microfluidic workflow for gene expression based profiling of drug combinations

4.1. Introduction

In this chapter, an expanded version of the introduced microfluidic and barcoding systems is described, that allows the multiplexed readout of gene expression from cells exposed to combinatorial droplet contents. Since the number of combinations increase exponentially with the number of drugs, conventional approaches for screening combinations can often exceed feasibility, especially when the biomaterial is limited (e.g. screening tumor biopsies). Additionally, various studies have shown the great potential of drug combinations for the treatment of cancer (Dietlein et al., 2015; Manchado et al., 2016) and bacterial (Brochado et al., 2018) or viral infections (Gulick et al., 1997). Together, it implies that the development of new approaches for combinatorial drug screenings can be beneficial. Droplet-based microfluidics would allow drug screenings in a miniaturized format, reducing sample volumes by several orders of magnitude compared to multi-titre plate approaches (Sackmann et al., 2014). Such a decrease in volumes can result in an increased throughput and scale since small amounts of reagents are handled and required. Additionally, a reduction of input material allows screenings with low cell numbers and, therefore, could enable testing drug combinations on tumor biopsies to define personalized treatments. Conventional drug screens focus on singular readouts (e.g. cell viability) to determine efficacies of drugs. These readouts do not provide insights into the molecular responses a drug or drug combination evokes in tumor cells. However, a deeper characterization of drug responses could be of great interest to decipher working principles of synergistic drug pairs and to determine potential causes of resistance towards drugs.

In order to perform combinatorial drug screenings in a miniaturized setting, the microfluidic workflow described in the previous chapter was improved in terms of scale and readout. By using the previously described valve-based system with an increased number of valves and pairing it with an autosampler-based system, we can produce hundreds of combinations in picolitre-sized droplets. In order to achieve a more comprehensive characterization of effects drug combinations have on cellular phenotypes, we used an advanced barcoding approach to perform highly multiplexed profiling gene expression.

4.2. Microfluidic workflow for a combinatorial drug screening in droplets

The generation of chemically complex droplets was achieved by combining the valve-based technology with a commercially available autosampler device (AS). A modular workflow with two microfluidic devices and an autosampler was established to enable rapid, automated and robust production of droplets with large numbers of compound combinations **(Fig. 4.1)**. (1) A valve-module with 24 inlets directed 20 injected compounds either to the waste outlets or the sample outlet. The two outermost valves on both sides of the device were used for oil injection to space out individual compounds with an immiscible phase. Opening of a single valve to generate a compound plug was followed by oil injection from all four oil valves **(Fig. 4.1A-1)**. The outlet of the valve-module was connected to a delay tubing long enough to be filled with one cycle of 20 compound plugs. Once the delay tubing was filled with all 20 compound plugs, two oil valves were opened to inject all compound plugs into the drop-maker chip. (2) The drop-maker combined compound plugs, cells and compounds injected from the autosampler into picolitre-sized droplets at a flow-focusing junction. The autosampler was used to aspirate one compound at a time from a 96-well plate, which was then combined with one cycle of 20 compound plugs coming from the valve-module **(Fig. 4.1A-2)**. Produced droplets containing a combination of two compounds and cells, were directly collected in a single tube. No washing steps were required as dead volumes and plug break-up at the inlet and outlet of the delay tubing were minimized (see below).

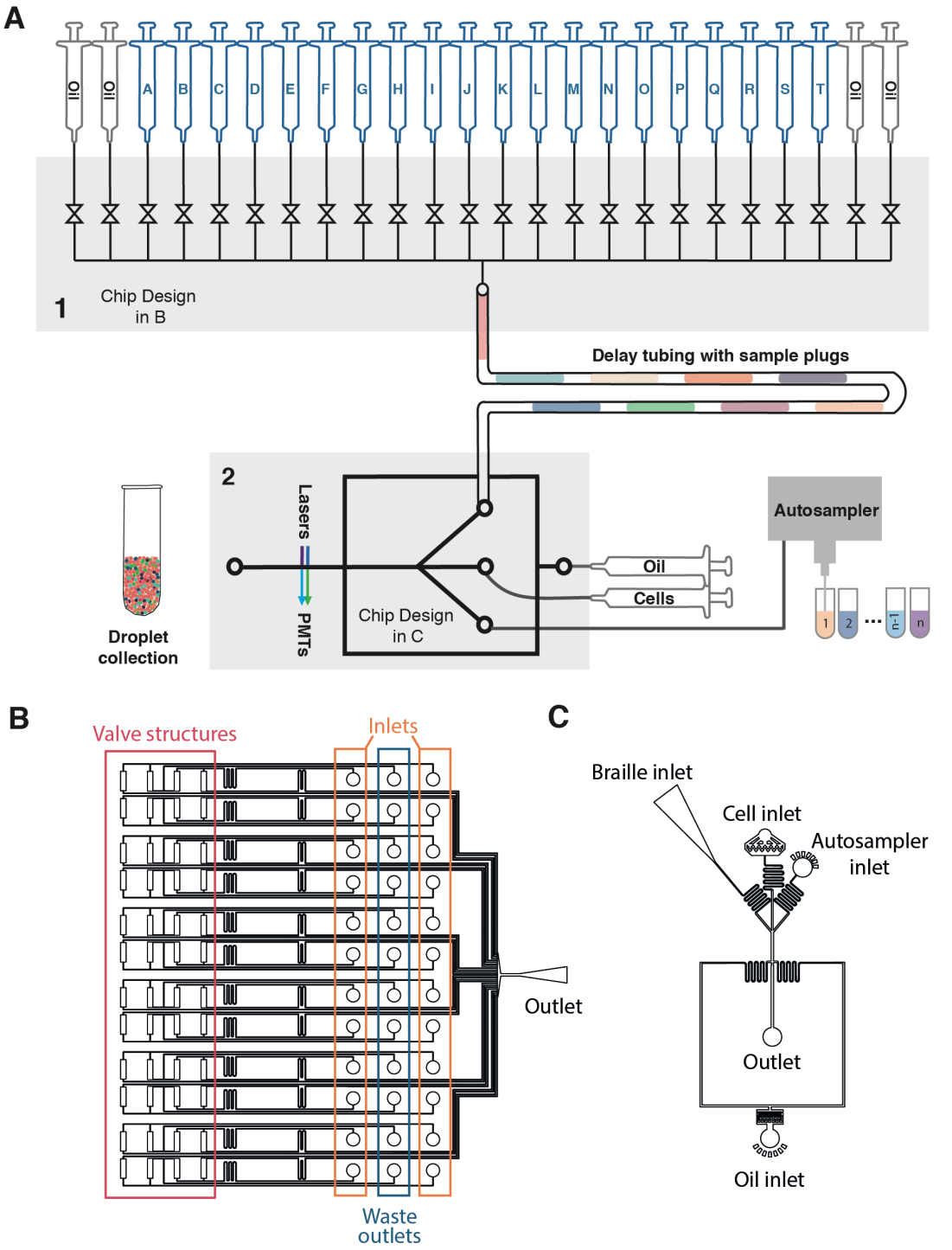

Fig. 4.1: Workflow for large-scale production of combinations in droplets

(A) Microfluidic pipeline used in this work to generate up to 420 drug combinations in droplets together with RNA-Seq based readout. The pipeline consisted of two microfluidic devices (B and C) and an autosampler. **(1)** A valve-module controlling 20 compound inlets and 4 oil inlets connected to syringes. Opening of one compound valve results in the injection of a particular drug into a delay tubing used to transfer a cycle of 20 compounds from the valve-module into the drop-maker. **(2)** On the drop-maker compound plugs were combined with cells and compounds from the autosampler before encapsulation into droplets, which were subsequently stored in a single tube. **(B)** Chip layout used for the valve-module with 24 inlets and valve units. **(C)** Chip layout of a drop-maker used to generate droplets from 3 aqueous inlets (compound plugs, cells and compounds coming from the autosampler).

For the valve-module, a microfluidic chip bound to an elastic membrane was aligned with its valve structures on pins of a braille display **(Fig. 4.1B)**. The outlet was a funnel-like structure that allows horizontal insertion of the delay tubing. The delay tubing, on the other end, was connected via a funnel-like inlet to the drop-maker device **(Fig. 4.1C)**. The drop maker was bound to an elastic membrane and glass slide enabling the side-wards insertion of the delay tubing **(Fig. 4.2A)**.

Determining the level of cross-contaminations between compound plugs

In order to avoid washing steps between individual compound plugs, the injection of compound plugs into the drop-maker was performed over a horizontal inlet. As compared to conventional vertical inlets, horizontal connection ports prevent cross-contamination between injected compound plugs due to plug breakup in dead volumes of the punched inlet (Clausell-Tormos et al., 2010). We injected compound plugs containing trypan blue dye to visualize the complete injection of plugs **(Fig. 4.2A)**. The individual frames of a video sequence showed that the entire compound plug was injected into the channel of the drop-maker. To further assess the purity of compound plugs injected for droplet generation, we supplemented syringes with Cascade blue or fluorescein and injected the compound dye mixtures in an alternating fashion (blue -> green -> blue etc.) **(Fig. 4.2B)**. The representative cycle of compound plugs illustrates that the level of cross-contaminations between plugs was below the detection limit as the fluorescence signal of a positive plug dropped to background level in the subsequent plug (one valve was not working resulting in 19 peaks). The level of cross-contaminations over all compound plugs (n=418) were analysed by comparing the median intensities of positive peaks for one dye (e.g. green positive peaks) and the median intensities of this dye in the negative peaks (e.g. green signal in blue peaks) **(Fig. 4.2C)**. Over all 22 cycles, we only observed cross contamination in four compound plugs. The summarized median green and blue intensities of green and blue peaks confirmed the low level of cross-contamination between compounds **(Fig. 4.2D, E)**. The level of cross-contamination was determined by comparing the median intensity of negative peaks to positive peaks. The overall cross-contamination of blue (by the green dye) and green (by the blue dye) peaks was 0.72% and 2.32%, respectively.

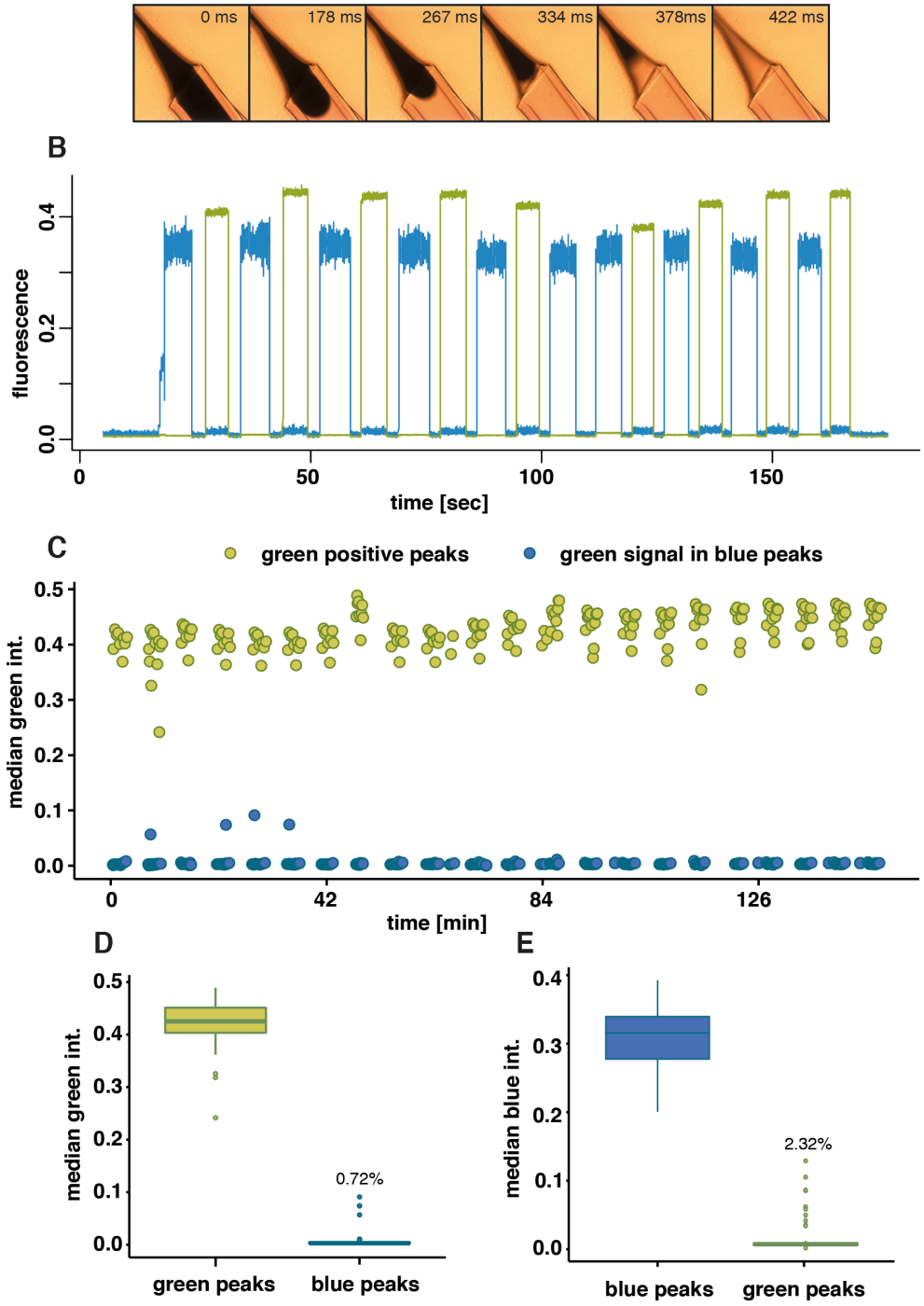

Fig. 4.2: Generation of compound plugs from the valve-module

(A) Horizontal injection of compound plugs from the valve-module into the inlet of the drop-maker. **(B)** Representative cycle of compound plugs injected into the drop-maker, each spaced out with oil. Compounds were supplemented with either green or blue dye and produced in an alternating fashion. **(C)** Median green intensities of green positive plugs and subsequent negative (blue) plugs over 22 injection cycles. **(D)** Summarized median green intensities from all green and blue peaks. Percentage of cross-contamination was calculated by comparing the median of green intensity of negative (blue) peaks with positive (green) peaks (0.72%). **(E)** Cross contamination of the blue dye in the green plugs. Median blue intensities of blue peaks and green peaks with the percentage of cross-contamination given for green peaks (2.32%).

4.3. Principle of generating chemical complexity in droplets

With the aforementioned approach, we produced combinations of compounds from the valve-module with compounds from a 96-well plate and encapsulated them into picolitre-sized droplets. Here a sequence of 20 compound plugs was combined with one compound injected by the autosampler, before repeating the cycle with the next compound coming from the autosampler. The workflow started with injection of plugs containing a single compound spaced out by an oil phase into a delay tubing **(Fig. 4.3A-1)**. The length of the delay tubing was adjusted to the volume of all 20 compound plugs plus oil spacer so that one entire cycle of 20 plugs was stored in it. The arrival of the compound plugs from the valve-module via the delay tubing and the compound from the autosampler at the drop maker needed to be synchronized. Therefore, after the first 5 compound plugs were injected into the delay tubing by the valve-module, the autosampler-based injection of a compound was started, since it took about 2 min until the compound reached its peak concentration in the drop-maker. During this time, all remaining compound plugs were injected into the delay tubing **(Fig. 4.3A-2)**. Once all plugs were produced, the two outer most oil valves from both sides of the valve-module were opened and all plugs were injected into the drop-maker, guaranteeing a stable and uniform flow rate for all compounds during their encapsulation into droplets. Each cycle of 20 compound plugs was combined with one compound injected from a 96-well plate by the autosampler **(Fig. 4.3B)**. We matched the injection time for each compound by the autosampler to the overall time of one cycle from the valve-module (generation and injection of compound plugs), thereby ensuring that all 20 compound plugs were combined with the specific compound injected from the autosampler. For cell-based screenings, we injected a single-cell suspension into the drop-maker at a concentration resulting in a single cell being encapsulated in approximately every fourth droplet (Clausell-Tormos et al., 2008). By combining several cycles of 20 compound plugs with compounds from the autosampler, we were able to produce high numbers of combinations at high throughput **(Fig. 4.3C)**. The overall time depends on the number of droplets (i.e. number of cells) screened for each combination and the number of compounds injected by the autosampler. In this study we combined 21 compounds from a 96-well plate and aimed for 1000 cells per combination. In this case the time for one combination was 16 sec and the total time for 420 combinations was 112 min.

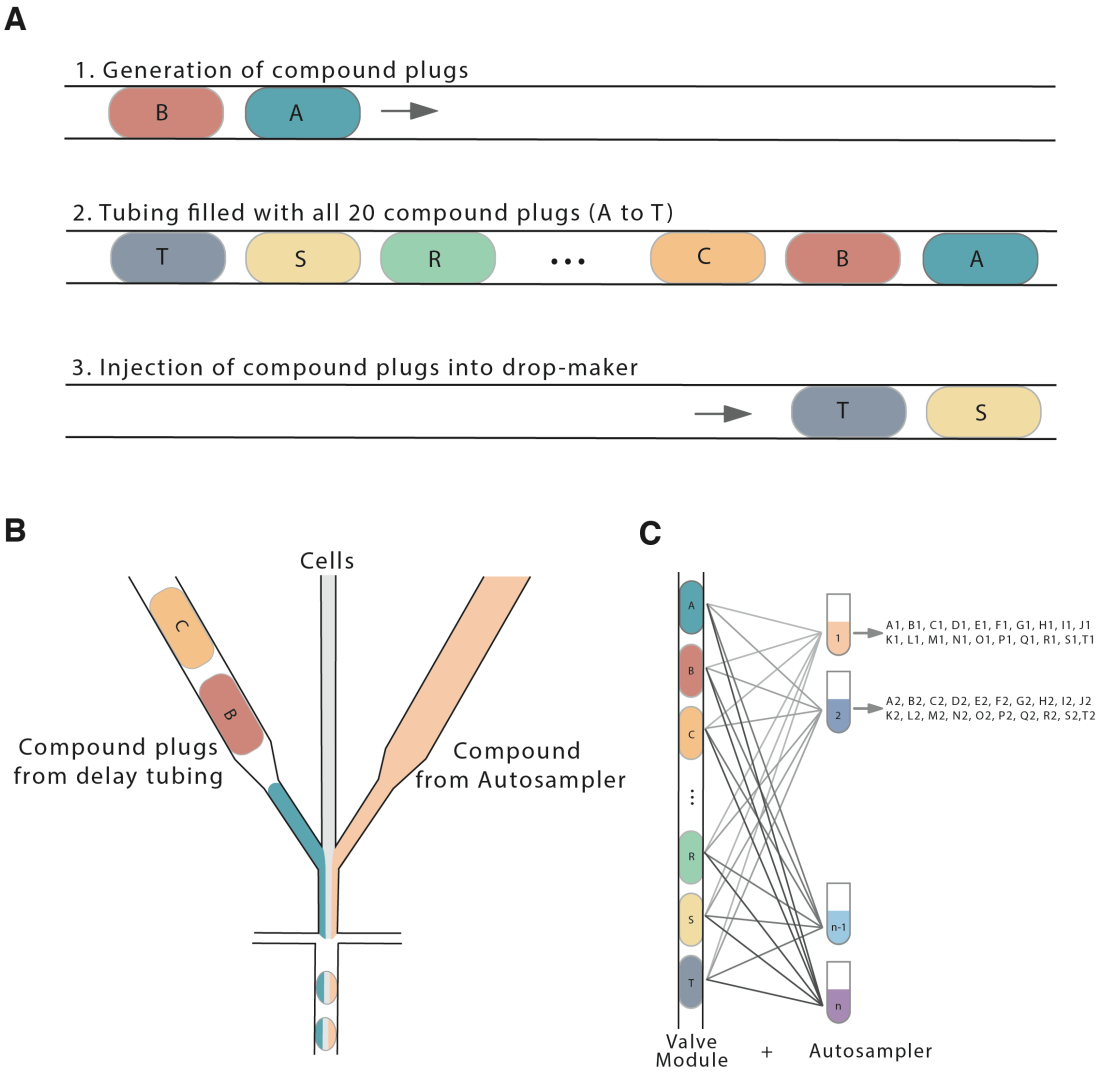

Fig. 4.3: Principle of combining compound plugs and compounds from the autosampler in droplets

(A) Plugs containing a single compound were produced by the valve-module and injected into a delay tubing, each spaced out by oil (1). The length of the delay tubing was adjusted to fit all 20 compound plugs. (2) Once all 20 compound plugs were produced and stored in the delay tubing (3), two oil valves were opened to inject all plugs into the drop maker. **(B)** Injection of compound plugs into a drop-maker along with cells and compound from the autosampler. All three streams formed a laminar flow before they were encapsulated together into droplets at constant ratios, giving rise to 20 combinations in a single cycle. **(C)** Compound plugs combined with 1 to n compounds from the autosampler results in a chemical complexity of 20n.

4.4. Combining compounds from the autosampler with sample plugs yields highly complex droplet libraries

The described principle of generating combinations from the valve-module and an autosampler was validated using fluorescence dyes. The autosampler was first loaded with a 96-well plate containing fluorescein and Cascade blue samples in an

alternating fashion, which were injected into a drop-maker device. An injection time of 5 min for each round of injection was set and fluorescence intensities were measured downstream of the inlet **(Fig. 4.4A)**. We observed clear separation of green and blue peaks with low dispersion into the carrier fluid (PBS). The overall cycle for one injection round was 6 min (1 min for the aspiration of one sample and 1 min for the signal to reach its plateau/max. concentration) leaving a time window of ~4 min during which one sample coming from the autosampler can be combined with all 20 compound plugs from the valve-module. Testing the performance of the pipeline to generate combinations in the described manner, compound plugs of Cascade blue were combined with fluorescein injected by the autosampler on the drop-maker device and co-encapsulated into droplets. Successful mixing was confirmed by measuring double-positive fluorescence signals of droplets from 9 cycles **(Fig. 4.4B)**. We observed one main population indicating that both compounds were encapsulated at constant ratios. This was confirmed by the median intensities of individual combinations, which were stable over all combinations with a CV of 2.9% and 3% for blue and green intensities, respectively **(Fig. 4.4E)**. The blue fluorescence intensity distributions of droplets from 9 cycles strongly overlapped and thus further indicated a stable co-encapsulation of the two dyes **(Fig. 4.4C)**. However, the blue fluorescence distributions were relatively broad and showed a high number of outliers **(Fig. 4.4C, E)**. To rule out that the observed variations in fluorescence intensities were caused by instable laminar flow regimes due to the oil spacer between each compound plug, we compared the density distributions of blue fluorescence intensities from droplets generated with or without oil spacer **(Fig. 4.4D)**. Since the observed variations of both distributions were comparable (CV spacer 13 % vs. CV no spacer 15%), we concluded that the variation did not derive from re-equilibrating flows between each sample plug but is likely to be a measurement artefact. Overall, this data demonstrated that mixing of compound plugs together with compounds from a 96-well plate into droplets can be achieved at constant ratios and thus this pipeline presents a novel approach for the fast and large-scale production of chemically diverse droplet libraries.

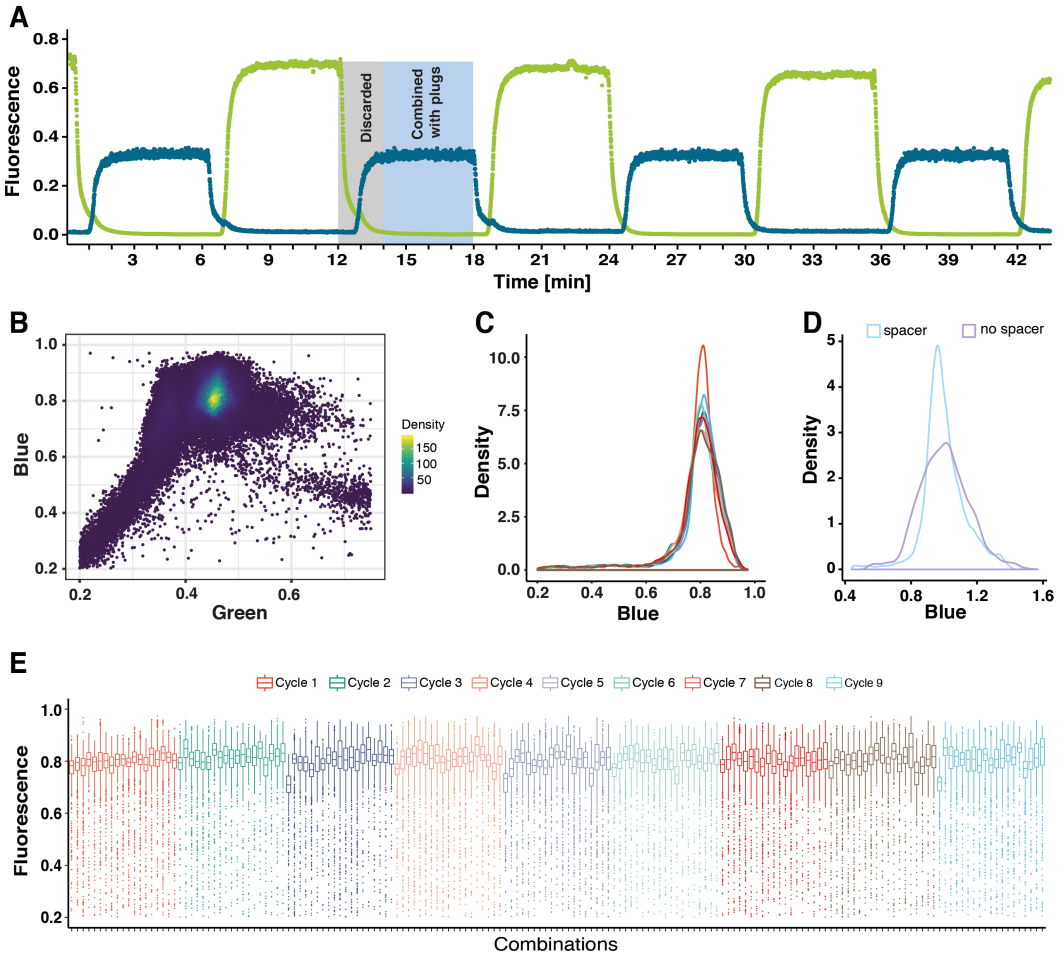

Fig. 4.4: Validations of large-scale production of combinations in droplets

(A) Fluorescence intensities of compounds injected by the autosampler labelled alternatingly with fluorescein and Cascade blue. Grey box represents time window during which no combinations were produced with compound plugs. Blue box indicates time window during which combinations made of compound plugs and autosampler compound can be generated. **(B)** Scatterplot of droplets produced from compound plugs labelled with Cascade blue and autosampler compounds labelled with fluorescein. Droplets from 180 combinations (9 cycles) were plotted. **(C)** Distributions of blue intensities from droplets over 9 cycles (180 combinations). Colour code for cycles identical to colour code in E. **(D)** Mean normalized distributions of blue fluorescence intensities from droplets generated with oil spacer or without oil spacer. **(E)** Blue fluorescence intensities of all individual combinations, colour coded according to cycles with each box corresponding to droplets generated from one combination.

4.5. Barcoding approach for gene expression based profiling of compound combinations

In order to deconvolve diverse droplet contents we implemented a combinatorial barcoding approach that barcodes the transcriptome according to the compounds a cell was exposed to. Thereby, we can achieve the simultaneous detection of treatment conditions (i.e. droplet contents) and their effects on gene expression by RNA-Seq. In this approach, two barcode species were joined in a ligation reaction to form one functional barcode **(Fig. 4.5A)**. BC-1 to BC-22 (BC-bio) were biotinylated and had a common primer site followed by a 10 bp barcode and a single-stranded ligation site. BC-A to BC-S (BC-dT) consisted of a complementary ligation site, a 10 bp barcode and a poly-dT sequence. The ligation of barcodes took place within droplets together in parallel with cell lysis and mRNA release. These mRNA molecules then hybridized to the polydT sequences of BC-dT. After ligation and hybridization, droplet contents were released and purified via the biotinylation site of BC-bio capturing only mRNA hybridized to ligated barcode combinations. Purified mRNA was reverse transcribed to cDNA and amplified using the common primer site on BC-bio and a template-switching oligonucleotide (TSO) as described in the Smart-Seq2 protocol (Picelli et al., 2013). Since only mRNA bound to combinations of two barcodes was purified and amplified, the barcoding approach was highly dependent on ligation **(Fig. 4.5B)**. This ensured depletion of mRNA with single barcodes and prevented formation of barcode combination during subsequent library preparation steps. We implemented the combinatorial barcoding approach for RNA-Seq based read-outs with the microfluidic pipeline by injecting drugs with unique barcodes (e.g. compounds injected into the valve-module were encoded by BC-dT (A-T) and 21 compounds injected by the autosampler from 96-well plates were encoded by BC-1 to BC-21 (BC-22 was used for separate untreated controls). Thereby, combinations generated by mixing compounds from the valve-module and the autosampler were encoded by two barcodes that were ligated to one functional barcode in droplets.

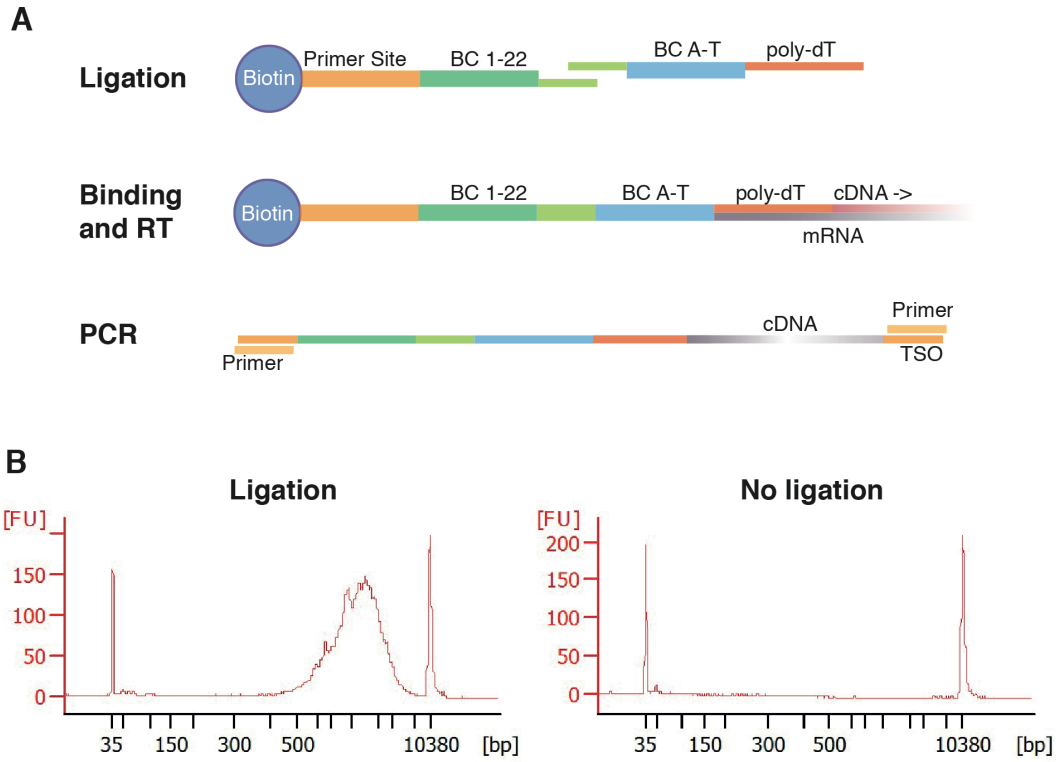

Fig. 4.5: Combinatorial barcoding of transcriptomes

(A) Deterministic barcoding of transcriptomes by a combination of two barcode species. Biotinylated barcodes (BC 1 -22) with a terminal primer site are ligated with barcodes having a 3'-end poly-dT sequence (BC A-T). Hybridized mRNA is isolated by streptavidin-coated beads and reverse transcribed (RT) into cDNA. Template switching oligonucleotides (TSOs) introduce a 5'-end priming site allowing whole transcriptome amplification. **(D)** Fragment length analysis of barcoded transcriptomes after amplification. The observed average peak size in the ligated sample was 2000 bp.

4.6. Combinatorial barcoding of cells in droplets for multiplexed RNA-Seq experiments

For RNA-Seq based read-outs of compound combinations in droplets, 20 compounds supplemented with the assigned barcodes were injected into the valve-module and combined with 21 barcode-compound mixtures injected from the autosampler. Both were encapsulated together with K562 cells into picolitre-sized droplets **(Fig. 4.6A-1)**. Since each droplet corresponds to one reaction vessel, uniquely barcoded according to its content by a set of two barcodes, we were able to pool all droplets in a single tube. After incubation, droplets were re-injected into a microfluidic device for pico-injection to add reagents for cell lysis and barcode ligation into each droplet **(Fig. 4.6A-2)**. Subsequently, the sets of two barcodes in

each droplet were ligated to form one functional barcode, and mRNA released from cells was hybridized to their poly-dT tails. Only after ligation of barcodes and hybridization of mRNA to the barcodes, droplets were broken and released droplet contents were used for preparation of the sequencing library **(Fig. 4.6A-3)**. The libraries comprised cDNA barcoded according to the droplet contents cells were exposed to, which in this case were 399 distinct treatment conditions (one valve was not working). After paired-end sequencing of barcodes and mRNA, reads were demultiplexed based on their barcode combinations, each encoding a specific droplet condition, and subsequently aligned. During demultiplexing 6 mismatches per barcode combination were allowed. We observed a high level of correlation of read counts per gene ($R^2 \geq 0.94$) between three replicates, implying good reproducibility of combinatorial barcoding of cells in droplets generated with the described approach **(Fig. 4.6B)**. Out of all reads assigned to the three replicates ($2.5x10^8$), 78% of reads were unambiguously demultiplexed to barcode combinations showing a low level of read losses. However, between individual samples (i.e. barcode combinations) we observed a wide range of read counts (2199 to 864845) and number of detected genes with more than one read (42 to 5092) **(Fig. 4.6C and Fig. 4.6D)**. To test whether the number of detected genes was sample (i.e. barcode) dependent, hierarchical clustering of scaled gene counts (Z-scores) was performed **(Fig. 4.6E)**. We found two main clusters based on barcodes from the valve-module (BC-A to BC-T) and one outlier cluster of samples with BC-O reflecting the low number of genes (42 to 566 genes) detected for these samples. We expected the number of genes detected per sample to be associated with the number of droplets generated per sample. However, we did not find correlations between the lengths of the compound plugs, which determined the number of droplets generated per sample, and the number of genes detected. Thus, the variation in gene counts is likely to be caused by varying demultiplexing efficiency of the barcode combinations or amplification biases during library preparation.

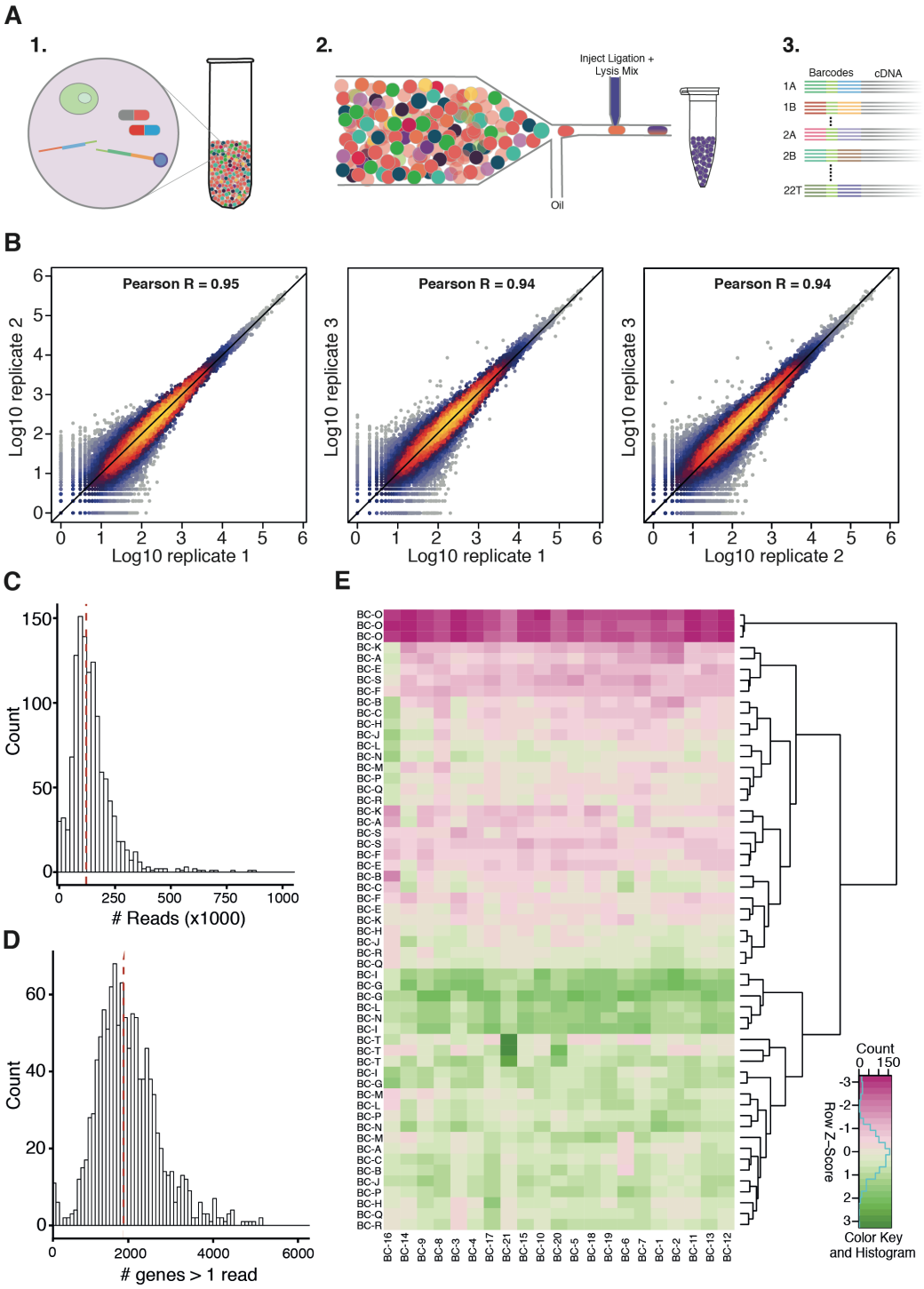

Fig. 4.6 Combinatorial barcoding for highly multiplexed RNA-Seq

(A) Illustration of the workflow used for combinatorial barcoding of transcriptomes. **(1)** Single cells were encapsulated into droplets along with a set of two barcodes and the corresponding pair of drugs. All droplets were stored in a single tube and after incubation droplets **(2)** were injected into a device for pico-injection. Reagents for cell lysis and barcode ligation were added to each droplet. **(3)** RNA-Seq libraries from cDNA with barcode combinations were prepared. **(B)** Pairwise correlations of read-counts per gene between three replicates after demultiplexing of reads according to their barcodes. **(C)** Read counts for individual samples (i.e. barcode combinations); the red line indicates the median read count (126000). **(D)** Counts of genes for individual samples covered by more than one read. The red line indicates the median gene count (1863). **(E)** Heatmap and hierarchical clustering of scaled gene counts. Each box represents a combination of two barcodes.

4.7. Discussion

The introduced workflow used a valve-based system and an autosampler to produce high numbers of drug combinations in picolitre sized droplets. Starting from the setup described in the first chapter, we significantly increased the number of generated combinations, as well as the throughput of the pipeline, as the washing steps became unnecessary. Identification of each combination was achieved by the deterministic barcoding of transcriptomes from cells encapsulated with a particular combination. Thereby, highly multiplexed gene expression experiments of compound combinations can be performed using droplet-based microfluidics.

4.7.1. Validation of compound plug purity

We designed a microfluidic device for the valve-module that prevented cross-contamination between compounds on the device by avoiding dead volumes. Additionally, horizontal ports were used to connect the delay tubing with the outlet of the valve-module and the inlet of the drop-maker to reduce plug-breakup at inlets due to dead volumes. By measuring compound plugs alternately labelled with a green or a blue dye, we could show that for the majority of compound plugs, no contamination was detected. The median intensity of green signals in blue peaks was reduced to 0.72% compared to green positive plugs, whereas the blue signals in green positive peaks was only reduced to 2.32%. This discrepancy in the contamination levels could arise from the fact that we observed more cross-contaminating green peaks in blue positive plugs (8 peaks in total) than blue positive peaks in green plugs (4 peaks in total). We also observed higher cross-contamination at the beginning of contaminated plugs, that declined exponentially and only caused contamination of the first generated droplets. This was caused by plugs breaking while entering the drop-maker device, and the subsequent plug picking up the residual liquids. By making use of the previously described sample collector (Section 3.5), we could prevent collection of the first droplets by setting the delay time for collection accordingly. However, this would need further optimization of the sample collector, since it was not applicable to high flow rates and caused high backpressures in the delay tubing.

4.7.2. Generating combinations from valve-module and autosampler based compound injections

In order to generate high numbers of combination we made use of autosampler-based injection of compounds from 96-well plates into the drop maker chip. By combining one compound from the autosampler with 20 compounds from the valve-module, it became possible to generate 20 different combinations, which could be increased multiplicatively with the number of compounds from the autosampler. We used a delay tubing for the valve-module based injection of compounds into the drop-maker. This setup has the advantage that one cycle of 20 compounds is first produced and then injected at a uniform flow rate since the same two oil valves are continuously open for the time of plug injection. As a result, no valves are switched during droplet production ensuring a stable droplet formation over time. Due to waiting times between compounds injected from the autosampler (~2min), the increase in time of using a delay tubing was only ~20 sec as compared to when the delay tubing is not used. The compound injection by the autosampler resulted in uniform peaks as shown by the fluorescence data. Each compound was combined with compound plugs to form droplets containing combinations. It was important to validate that compound plugs and compounds from the autosampler were encapsulated at a constant ratio in order to ensure that both compounds were encapsulated at the right concentration. The validation of the pairing between blue positive compound plugs and green positive compounds from the autosampler showed that one major double positive droplet population was generated (Fig. 4.4B). We found the median intensities over all generated combinations to be stable with a CV of 2.9% and 3% for the blue and green signal, respectively (Fig. 4.4E). This indicates that both compounds were encapsulated into droplets at a highly constant ratio. Since both dyes were encapsulated at a constant ratio, the concentrations of compounds encapsulated into droplets correspond to injected concentrations diluted by a factor given by the flow rate ratios between all injected aqueous phases (valve-module: 0.5; autosampler: 0.25; cells: 0.25). However, the observed intensity distributions over all cycles were broad indicating variations between individual droplets. We anticipated that this observation might be explained by unstable laminar flows of the two dyes caused by the oil spacers between each compound plug. However we found the same level of variation in fluorescence intensities when measuring droplets produced without oil spacers

(Fig. 4.4D). Therefore, we concluded that the variation of fluorescence intensities around a constant median intensity was caused by recoding droplets generated from a laminar flow of three aqueous phases and the fluctuation of the droplet trajectory within the channel (as discussed in 3.8.1). Taken together, the discussed results of the presented microfluidic workflow is applicable to generate up to 420 chemically distinct combinations, each present in around 2500 droplets, in less than 2h.

4.7.3. Barcoding of transcriptomes according to droplet contents

To characterize effects on gene expression caused by a combination of two compounds, we established a barcoding approach in which each cells transcriptome was labelled according to the compounds it was exposed to within a droplet. Ligation of two barcodes was used to encode a combination but additionally mRNA was hybridized to the barcodes and allowed its reverse transcription. Consequently, it became possible to barcode transcriptomes from cells according to combinatorial drug treatments, using a combination of two barcodes. The designed approach was highly dependent on the ligation of two barcode species since both were crucial for later library preparations, ensuring an RNA-Seq library containing only cDNA fragments with two barcodes. Additionally, biotin-streptavidin based purification was used to avoid PCR-based joining of non-ligated fragments, as seen by the lack of amplified fragments in the no ligation control (Fig. 4.5B). Thus, the approach can be used to perform deterministic barcoding of cells in droplets, since we assigned each barcode to a specific treatment condition. This is in contrast to widely applied random barcoding in droplets used for single cell sequencing (Klein et al., 2015; Macosko et al., 2015). Here, thousands of droplets (i.e. cells) are uniquely barcoded, however, without knowing the barcode sequence. Our approach does not aim at single barcodes for each cell, but barcoding each condition with a unique set of barcodes. In a screen we combined 19 compound-barcode mixtures from the valve-module (one valve was not working) with 21 compound-barcode mixtures from the autosampler, resulting in a total of 399 combinations of barcodes (i.e. drugs). The integration of barcoding mRNA from cells encapsulated into droplets, made it necessary to perform pico-injection to add reagents for cell lysis and ligation. Although each droplet was processed individually in this step, high frequencies (~300 droplets/sec) allowed processing all droplets in approx. 1h,

which reduced biases in the incubation times over the course of pico-injections. Within each droplet, barcodes were ligated and released mRNA hybridized to their poly-dT tail. Since the emulsion was broken in a volume much larger than the volume of droplets (a ~$3.125x10^7$ fold increase), the chance of unbound mRNA hybridizing randomly to barcodes was reduced ensuring a high level of specific barcoding. However, to determine the level of cross-hybridized mRNA, artificial mRNA species could be spiked into the supernatant. The pairwise correlation between read counts per gene for all replicates was comparable to described correlations of 3'-end RNA-Seq libraries (Hennig et al., 2018), demonstrating a high degree of reproducibility between replicates. Although, the read loss due to demultiplexing reads according to their barcodes (78% of reads were assigned to barcode combinations) was low, we observed only a median of 1863 genes covered by more than one read. The total amount of reads for all samples from three replicates was $2.5x10^8$, leaving only approx. $2x10^5$ reads for each sample. This is a sequencing depth generally considered for high-throughput single cell sequencing methods (Hwang et al., 2018) and thus, detections of ~2000 genes per sample are comparable to counts expected from single cell sequencing experiments (Macosko et al., 2015; Torre et al., 2018). This indicates towards a comparable sequencing efficiency of the presented combinatorial barcoding approach. An increase in the number of genes per sample can be achieved by increasing the sequencing depth as indicated by samples sequenced at a higher coverage (average of $7.7x10^6$ reads), which resulted on average in 12216 detected genes per sample. We analysed if variations in gene counts are due to different amounts of droplets generated per combination. However, the length of compound plugs failed to explain the number of detected genes since no correlation between plug length and gene counts was observed. Additionally, cluster analysis of gene counts did not reveal barcode dependent clustering for the majority of barcodes. We observed two main clusters based on BC-A to BC-T, which again did not show barcode specificity. Therefore, we concluded that the observed variation was likely due to biases in the amplification efficiency between different fragments. Furthermore, differences in the efficiency of demultiplexing reads on barcode combinations could explain the observed variations in gene counts per sample. Taken together, deterministic barcoding of combinations was applicable for the demultiplexing of RNA-Seq data according to combinatorial drug treatments in droplets.

5. Highly multiplexed gene expression profiling of drug combinations in droplets

5.1. Introduction

In the following chapter we applied the microfluidic pipeline introduced in chapter 4 to perform combinatorial screens using ant-tumor drugs in droplets and analysed their effects on gene expression in K562 cells. Gene expression based profiling of drugs was proven to be powerful in determining the mode of action and repositioning of drugs (Subramanian et al., 2017; Woo et al., 2015). As compared to singular readouts in conventional high-throughput drug screens, such as apoptosis or enzyme activity, gene expression of cells under perturbation enables more comprehensive readouts of the drugs' effect on cells, which can be useful to improve models aiming at predicting drug response or understanding of drug resistance and sensitivity. Since RNA-Seq experiments are expensive, several low cost approaches have been developed to combine high-throughput drug screens with gene expression profiling. Increased sequencing throughput was achieved by focusing only on gene panels (Subramanian et al., 2017) or by multiplexed RNA-Seq experiments (Bush et al., 2017; Ye et al., 2018). However, these approaches aim at studying the effects of single drugs on gene expression and screens were still performed in plate-based setups. The translation of drug screens from plates into droplets harbours the potential to further increase the multiplexing capacity of RNA-Seq based analysis of drug responses. By applying the described microfluidic pipeline for gene expression based profiling of drug combinations, their effect on the transcriptome can be studied in a highly multiplexed manner. Gene expression data from tumor cells treated with drug combinations can be used to determine drug synergy based cell death signatures (Szalai et al., 2018). Subsequently, the datasets can be used to decipher the underlying mechanisms of the observed

synergism based on perturbation responses in gene expression (Schubert et al., 2018). Furthermore, using droplets results in a strong reduction of sample volume and allows screening high numbers of drug combinations (420) with low numbers of cells (~$2x10^6$). Low amounts of input material can facilitate drug screens on tumor biopsies to determine personalized treatments (Eduati et al., 2018). Additionally, encapsulating single cancer cell together with drug combinations into droplets provides a potential platform for single cell characterization of drug sensitivity, and thus, to study the effect of tumor heterogeneity on drug response using gene expression profiling.

5.2. Drug library used for RNA-Seq based screens

Screening drugs in surfactant-stabilized droplets can potentially result in micelle-mediated exchange of drugs between droplets (Gruner et al., 2016). The distribution (LogD value) or partitioning coefficient (LogP value) can be used as a measure to assess the probability of a drugs' retention in a water droplet surrounded by oil. A compound with a negative LogD-value has a higher solubility in water than in octanol and thus has a higher probability to be retained in a water droplet. Therefore, we selected drugs based on two criteria: First, drugs from the ChEMBL database were selected based on a low LogD value and then drugs were further selected by their potential relevance for cancer **(Tab. 5.1)**. The majority of the selected 40 drugs have negative LogD-values or smaller than one, apart from Imatinib and Trametinib (2.49 and 2.54). Targets of the selected drugs were used to generate a network based on protein-protein associations using STRING **(Fig. 5.1A)**. We obtained a dense network with strongest enrichments for cancer related, PI3K-Akt and Ras signalling pathways (FDR of 2.6e-24, 3.68e-24 and 1.29e-22, respectively). To determine the sensitivity of K562 cells towards the selected drugs, the growth reductions after 48h over hundred fold concentration ranges were measured. Growth reduction curves were used to calculate GR20 values (Hafner et al., 2016), which were used in all later drug screenings **(Fig. 5.1B)**. For drugs where no effect on the growth of K652 cells was measured, a concentration of 100 µM was used. Taken together, despite the restriction to drugs with low logD values, a cancer relevant drug library was selected out of which the majority of drugs reduced the growth of K562 cells. Therefore, the selected drug library was used to generate drug combinations exploiting the pipeline described in chapter 4.2.

Tab. 4.1: Drug library

Drug	Targets	LogD pH 7.4	GR20 [µM]
10Z-Hymenialdisine	MEK1	0.20	6.085
5-Fluorouracil	TYMS	-1.64	0.571
5-Iodotubercidine	ADK/INSR/ PKA/CK1	0.20	5.705
AT9283	AURKA/AURKB/JAK2/JAK3	-0.99	0.147
Baricitinib	JAK1/JAK2	0.46	0.387
Blebbistatin	MYH2	1.09	77.16
Clofarabine	RRM1 / DNA Polymerase	0.45	0.455
Cytarabine	DNA / RNA Polymerase	-0.81	100
Dacarbazine	DNA Animetabolite	0.03	1,867
Decitabine	DNMT1	-1.87	0.206
Dexrazoxane	TOP2	-0.91	11.842
Dovitinib	FLT3/c-Kit/FGFR1/FGFR3	0.80	0.034
Doxorubicin	TOP2	-1.47	0.322
Epirubicin	TOP2	-1.47	0.100
Fludarabine Phosphate	DNA Animetabolite	-7.62	6.737
Gemcitabine	DNA Animetabolite	-2.22	0.00292
Gimeracil	DPYD	-2.56	100
H-7 dihydrochloride	PKC/PKG/PKA	NA	64.89
Hematoxylin	EGFR/ERBB2/c-MET/c-KIT/SRC*	-0.18	0.625
Imatinib	BCR-ABL	2.49	0.039
Methodextrat	DHFR	-5.10	100
Mitomycin C	DNA Synthesis	-0.30	0.815
Nelarabine	DNA Animetabolite	-0.26	100
NMS-1286937	PLK1	-1.19	0.105
Olomoucine	CDK2/ERK1	-0.2	112.1
Oxaliplatin	DNA	-0.47*	8.699
PF-562271	FAK	-0.75	7.059
Pomalidomide	TNF-alpha	-0.71	100
Sangivamycin	PKC	-2.28	100
SB-747651	MSK1/MSK2	-2.24	0.285
SF-1126	PI3K/mTOR	-2.87**	0.257
Sonolisib	PI3K	-0.95	1.460
Streptozotocin	DNA	-1.25	100
Sunitinib Malate	VEGFR2/PDGFRb	0.44	6.702
Tabloid / Thioguanine	DNMT1	-1.50	0.813
Thiostrepton	FOXM1	0.60	5.932
Trametinib	MEK1/2	2.54	0.0977
Triciribine	AKT1/AKT2/AKT3	-7.22	1.992
Wortmannin	PI3K	-1.64	0.223
YM155	BIRC5	-1.45	0.00041

Targets according to the manufactures information
LogD pH 7.4: ACD_LogD pH 7.4 values at given by ChEMBL database using ACD software v12.01
* LogP value from DrugBank
** ALogP value ChEMBL: Atomic LogP value (Ghose and Crippen, 1987)

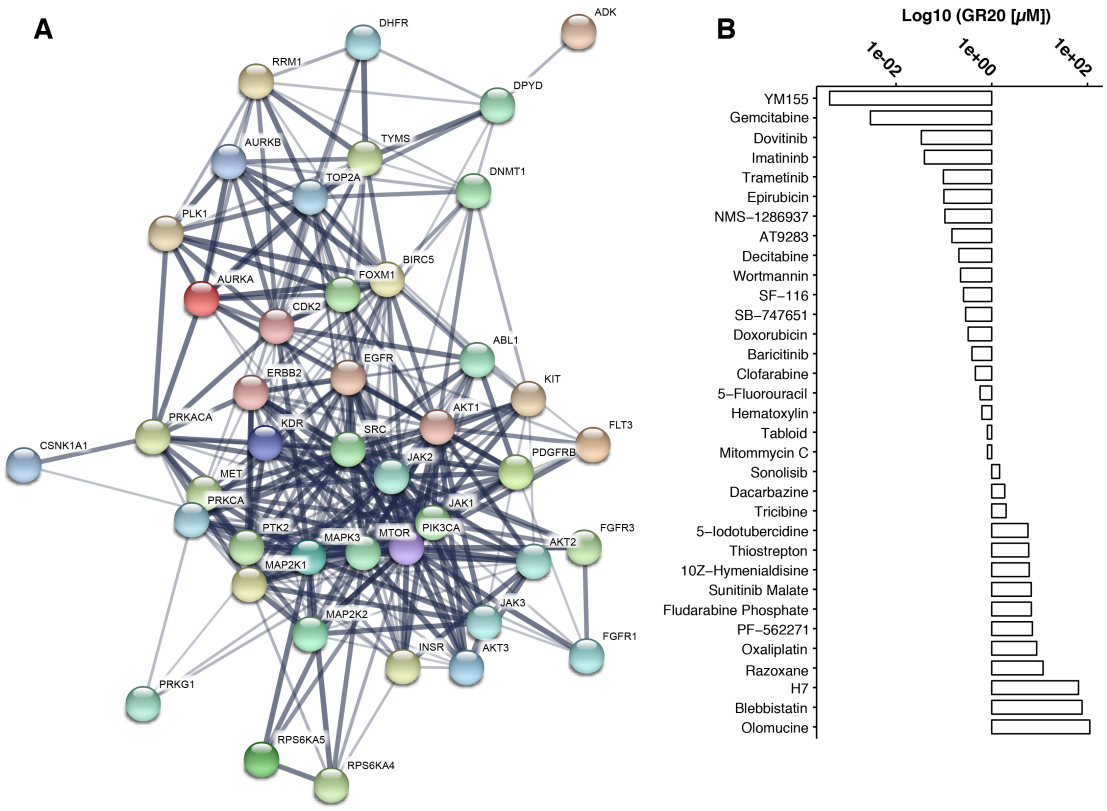

Fig. 5.1: Network analysis and growth reduction of drugs

(A) STRING-based network of associations between proteins targeted by the drugs used in this work. Each node represents all proteins that can be produced from a single protein-coding gene. Edges illustrate protein-protein association and their thickness indicates confidence of their association. Drug targets were strongly enriched for cancer related signalling pathways. **(B)** Log10 values of drug concentrations at which the growth of K562 cells was inhibited by 20% (GR_{20}) after an incubation of 48h.

5.3. Screening of drug combinations in droplets and detection of their effects on gene expression

For combinatorial drug screenings 19 drugs from the drug library (**Tab. 5.1**) and a DMSO sample were assigned to 20 BC-dTs (BC-A to BC-T) and mixtures were injected into the valve-module of the microfluidic pipeline described in chapter 4.2. A remaining set of 20 drugs and a DMSO sample were assigned to 21 biotinylated BCs (BC-1 to BC-21) and mixtures were aspirated from 96-well plates by the autosampler. Combinations of drug-barcode mixtures from the valve-module and autosampler were encapsulated into droplets along with K562 cells and incubated for 12h. For each replicate an additional emulsion of untreated control droplets (DMSO control) was prepared and incubated separately. Subsequently, the emulsion comprising 420 drug combinations was mixed with the DMSO control

emulsion and droplets were processed for library preparation according to the workflow described in chapter 4.7. To test potential biases based on the drug source (valve-module or autosampler) or barcode fragments used to encode drugs, we run two screens with 420 combinations generated from four sets of drugs: Out of the four sets always two sets were either used on the valve-module or the autosampler and either encoded with BC-dT or BC-bio barcodes, respectively **(Tab. 5.2)**. In screen 1 the drug sets 1 and 3 from the valve-module were combined with the drug sets 2 and 4 from the autosampler. In screen 2 the drug sets 1 and 4 from the valve-module were combined with the drug sets 2 and 3 from the autosampler. In total 840 drug combinations were generated in droplets out of which 630 were unique. Out of the 210 combinations that were generated in both screens, 110 combinations were produced in the same order and were encoded by the same barcode combinations (e.g. Drug-A BC-dT + Drug-B BC-bio and Drug-A BC-dT + Drug B BC-bio). 100 drug combinations were shared between screen 1 and 2, but generated in a different order, and thus, were encoded by the opposite order of barcodes (e.g. Drug-A BC-dT + Drug-B BC-bio and Drug-B BC-dT + Drug-A BC-bio).

Tab. 5.2: Drug combinations screen 1 and screen 2

Drug set 1	Drug set 2	Drug set 3	Drug set 4
Clofarabine	10Z-Hymenialdisine	5-Iodotubercidine	5-Fluorouracil
Cytarabine	Blebbistatin	Baricitinib	AT9283
Dovitinib	Gimeracil	Doxorubicin	Dacarbazin
Fludarabine Phosphate	Methodextrate	H7	Decitabine
Gemcitabine	Nelarabine	Hematoxylin	Epirubicin
Imatininb	Olomucine	PF-562271	Mitomycin
Oxaliplatin	DMSO	Razoxane	NMS-1286937
DMSO	SF-116	SB-747651	Tabloid
Pomalidomide	Sonolisib	Tricibine	Trametinib
Sunitinib Malate	Streptozotocin	Wortmannin	YM115
	Thiostrepton		

	Valve module BC dT	Autosampler BC bio	Number of combinations
Screen 1	Drug set 1 (10) Drug set 3 (10)	Drug set 2 (11) Drug set 4 (10)	10 x 11 = 110 → Same combo same order 10 x 10 = 100 → Unique 10 x 11 = 100 → Unique 10 x 10 = 100 → Same combo different order
Screen 2	Drug set 1 (10) Drug set 4 (10)	Drug set 2 (11) Drug set 3 (10)	10 x 11 = 110 → Same combo same order 10 x 10 = 100 → Unique 10 x 11 = 100 → Unique 10 x 10 = 100 → Same combo different order

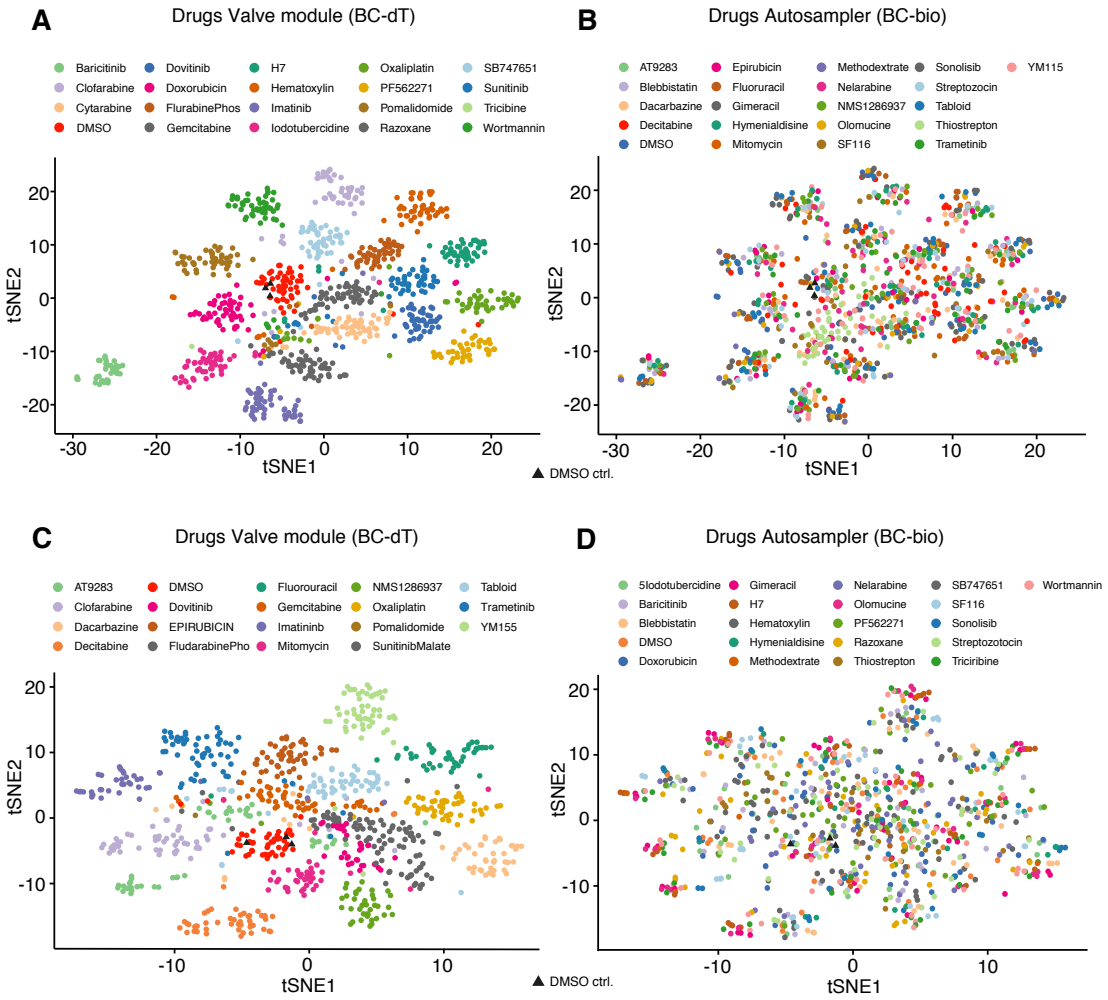

Fig. 5.2: Dimensional reduction of gene expression data from drug combinations

Normalized gene expression data of drug combinations from screen 1 (A and B) and 2 (C and D) was visualized using t-distributed stochastic neighbourhood embedding (tSNE). **(A)** Clustering of gene expression data from 420 treatment conditions generated in screen 1 coloured based on the 19 drugs and DMSO from the valve-module encoded by barcodes BC-dT. Each data point represents a treatment condition encoded by a barcode combination and triangular shaped data points refer to separately incubated DMSO controls. **(B)** Cluster analysis of drug combinations from (A) coloured according to the 20 drugs and DMSO injected by the autosampler. **(C)** tSNE-based clustering of gene expression data from 399 treatment conditions generated in screen 2. Clustered data points were colour coded according to drugs from the valve-module encoded by BC-dT with each data point representing a treatment condition. Triangular shaped data points refer to DMSO controls. **(D)** Clustering of gene expression data from screen 2 coloured according to 21 drugs from the autosampler encoded by BC-bio.

Sequenced libraries from three replicates per screen were demultiplexed according to their barcode combinations and gene expression profiles for all drug combinations were obtained. Data was preprocessed and normalized to account for differences in library size and batch effects. In order to illustrate expression profiles from all 420 drug combinations, we performed dimensional reduction using t-distributed stochastic neighbourhood embedding (tSNE) of the normalized gene

expression data **(Fig. 5.2)**. In both screens we observed strong clustering based on drugs from the valve-module, which were encoded by barcodes BC-A to BC-T **(Fig. 5.2 A, C)**. Apart from a few exceptions, we observed one main cluster for each drug injected from the valve-module encoded by a particular BC-dT. Additionally, all samples containing DMSO from the valve-module fall into one cluster that additionally contains the DMSO control samples. However, we did not observe any clustering based on drugs injected from the autosampler encoded by BC-bio (BC-1 to BC-22) in both screens **(Fig. 5.2 B, D)**. The dominant effect of drugs from the valve-module on gene expression was further confirmed by spearman correlations of drug signatures for screen 2 **(Fig. 5.3)**. We observed the highest degree of correlation between combinations of the same two drugs. However, the correlations between signatures where drugs from the valve-module were the same, showed much stronger correlations as compared to signatures of drugs from the autosampler. Since the degree of correlation was comparable to signatures between different drugs, we concluded that effects of drugs from the autosampler on gene expression were undetectable in the screens.

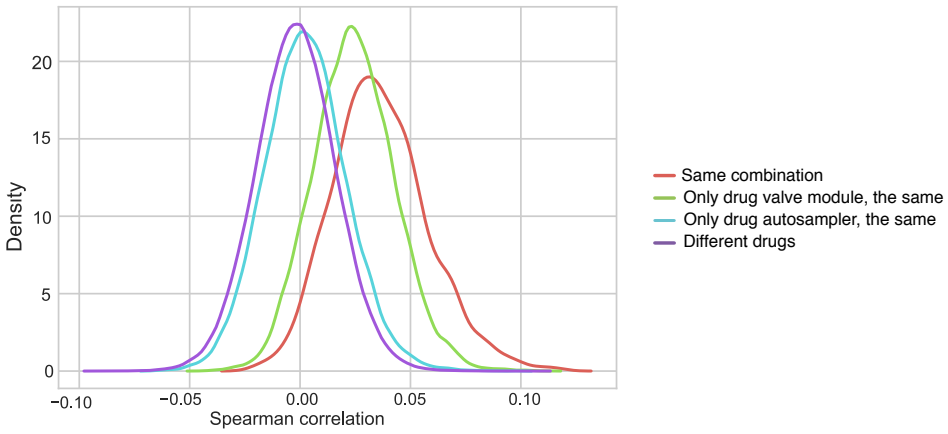

Fig. 5.3: Correlations between drug signatures

Distributions of spearman correlations between drug signatures from screen 2. Signatures were determined by z-scores of gene expression for each sample. Correlations between the same combinations (red, n=3), all combinations containing the same drug from the valve-module (green, n=21), all combinations containing the same drug from the autosampler (blue, n=19), and combinations containing different drugs (purple, n=378 or n=380). Analysis and plots by Bence Szalai (Julio Saez-Rodriguez Group; BioQuant, University Heidelberg).

5.4. Improved detection of drug combinations by diminishing residual non-ligated barcodes

Since the observed clustering was dependent on position and barcode (Valve-module and BC-dT), we excluded any biological relevant reason. We hypothesized that the barcoding protocol was most likely the cause of the observed dominant effects by drugs from the valve-module. In particular, non-ligated biotinylated barcodes (BC-bio) were purified on beads together with barcode combinations and mRNA. The common primer site on non-ligated BC-bio fragments caused their linear amplification during PCR and products were likely to function as primers on ligated barcode combinations. This would result in the loss of gene expression information encoded by BC-bio, since free-floating BC-bio fragments could randomly bind to barcoded cDNA, resulting in a PCR product with new barcode combinations. To avoid priming by non-ligated BC-bio fragments we improved the library preparation protocol by cleaving all fragments off the streptavidin coated beads and size selecting fragments longer than 300 bp. We prepared new libraries using the improved protocol from cDNA of screens 1 and 2 and re-sequenced the samples. Cluster analysis based on normalized gene expression data from screen 1 (see appendix Fig. 9.1) and 2, demonstrated an improvement since no dominant effects of drugs from the valve-module were observed anymore **(Fig. 5.4)**. Instead, no strong clustering based neither of drugs from the valve-module **(Fig. 5.4A)** nor of drugs from the autosampler **(Fig. 5.4B)** was observed. We again compared gene expression drug signatures from both screens by Spearman correlation and found the level of correlations for drugs from the valve-module and the autosampler to be comparable **(Fig. 5.4 C, D)**. Gene expression signatures of the same drug combinations showed in both cases the highest degree of correlations. This result implied an improved detection of drug effects by gene expression for drugs encoded by BC-bio in both screens. Finally, we performed cross-validations between combinations from screen 1 and 2. We compared signatures from the same drug combinations generated in the same order and encoded by the same barcode combination and observed a good level of correlation between drug signatures **(Fig. 5.4E)**. However, when comparing signatures of the same drug combinations generated in a different order (e.g. Drug A autosampler BC-bio with BC-B valve-module BC-dT), the observed correlation between drug signatures from the two screens was low.

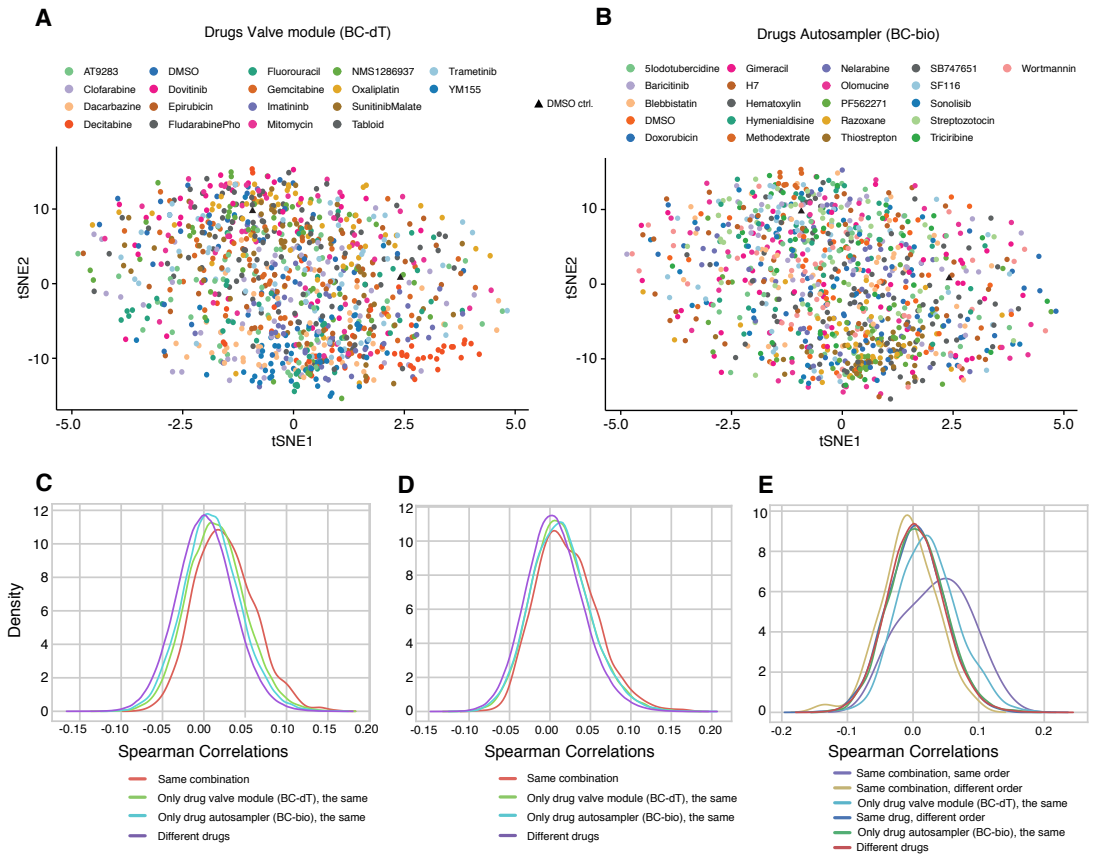

Fig. 5.4: Analysis of re-sequenced libraries after removal of non-ligated BC-bio.

(A) Clustering of gene expression data from 399 combinations generated in screen 2 coloured according to 17 drugs and DMSO from the valve-module encoded by barcodes BC-dT. Each data point represents a treatment condition encoded by a barcode combination **(B)** Cluster analysis of drug combinations from (A) coloured according to the 20 drugs and DMSO injected by the autosampler encoded by BC-bio. **(C)** Spearman correlations between drug signatures from screen 2, calculated from z-scores of gene expression in each sample. Signatures from combinations with the same drugs (red, n=3), combinations where only drugs from the valve-module (green, n=19) or the autosampler (blue, n=21) were the same and combinations with different set of drugs were compared **(D)** Spearman correlations between drug signatures for screen 1 as in (C). **(E)** Cross validation of drug signatures between screen 1 and 2 by spearman correlations. Correlations of gene expression profiles from the same drug combinations generated in the same order (purple curve; Drug A BC-dT + Drug B BC-bio with Drug A BC-dT + Drug B BC-bio), the same drug combinations generated in different order (yellow curve; Drug A BC-dT + Drug B BC-bio with Drug B BC-dT + Drug A BC-bio), the combinations containing the same drug form the valve-module (bright blue curve, Drug A BC-dT with Drug A BC-dT), the combinations containing the same drug form the autosampler (green curve, Drug A BC-bio with Drug A BC-bio), the combinations containing the same drug generated in different order (blue curve, Drug A BC-dT with Drug A BC-bio) and combinations containing two different drugs (red curve). Analysis and plots of spearman correlations by Bence Szalai (Julio Saez-Rodriguez Group; BioQuant, University Heidelberg).

Comparably, the correlations between signatures of the same drugs encoded by BC-dT in both screens were good, but correlations between the same drugs encoded by BC-dT in one screen and BC-bio in the other screen were as low as comparisons between different drugs. Therefore, we concluded that we still observed barcode

(i.e. injection position) dependent effects in the gene expression data. Nonetheless, removing residual non-ligated barcode fragments resulted in improved detection of effects from both drugs in individual experiments, indicating that successful barcoding of drug combination can be achieved.

5.5. Discussion

In the presented work we aimed at generating the first gene expression data set from cells perturbed with hundreds of drug combinations. We envisioned determining synergistic drug pairs from gene expression data based on cell death signatures. Additionally, analysing pathway activities in cells under perturbation would allow finding potential mechanisms of the observed synergistic effects. However, we did not detect robust effects of drug combinations in the collected gene expression data. Initially, we only observed strong clustering of the gene expression data based on one set of drugs (encoded by BCs-dT) but no clusters were formed based on perturbation with drugs from the autosampler (encoded by BCs-bio). The observed clusters in both screens demonstrate that the applied microfluidic approach was able to pick up drug specific effects, since the majority of the 20 treatment conditions in both screens form single defined clusters (Fig. 5.2). The distinct clustering observed for all 29 drugs from both screens indicate that no detectable responses based on drug exchange were observed. This conclusion was confirmed by the formation of one defined cluster of all DMSO samples generated from the valve-module and the DMSO control samples.

We hypothesized that the dominant effects of drugs encoded by BC-dT were caused by residual BC-bio functioning as primers during whole transcriptome amplifications. An optimized protocol for library preparation in which we managed to remove remaining non-ligated BC-bio from the cDNA libraries indeed resulted in improved sequencing results. Compared to the initial results we did not observe cluster formation in the tSNE analysis neither by drugs from the valve-module nor the drugs from the autosampler (Fig. 5.3A, B). The previously observed dominant effects of drugs encoded by BC-dT were undetectable or diminished for screen 1 or screen 2, respectively, as shown by correlations between drug signatures from drugs encoded by BC-dT or BC-bio (Fig. 5.3C, D). Correlations between combinations containing only one specific drug from the valve-module or only one specific drug from the autosampler are comparable, while correlations between drug signatures from the same combinations were increased. Comparison within the individual screens indicated towards a detection of both drugs. Since we exchanged half the drugs between screen 1 and 2, we could compare drug signatures of the same combinations across two experiments, either prepared in the same or in the opposite order. Order in this case refers to injection position

(valve-module or autosampler) and barcode species (BC-dT or BC-bio) used for demultiplexing gene expression data to each treatment condition. Since we observed only a high degree of correlations between drug signatures of the same combinations generated in the same order but not for the one generated in a different order, the data was still biased by the barcoding position. Despite the initial improvements, we could not fully overcome the positional bias in the gene expression data. However, removing residual non-ligated BC-bio diminished the strong clustering of drugs encoded by BC-dT and resulted in an improved correlation between signatures of drugs encoded with BC-bio in the individual experiments. Therefore, it is likely that the positional bias between drug signatures across two experiments was still caused by remaining non-ligated barcodes. Potential solutions offer more stringent purification protocols in order to ensure the complete removal of unbound and non-ligated barcodes. Additionally, increasing the sequencing depth could improve the detection of drug effects. For screen 1 and 2 each treatment conditions was sequenced approximately with a depth of $2.4x10^5$ and $2x10^5$ reads, respectively. Considering that these are the target sequencing depth for single cell sequencing experiment (Hwang et al., 2018), it is not surprising that the number of detected genes (genes > 1 read) are comparably low (median of 1863 genes per sample in screen 2) as for droplet based single cell sequencing methods (Zhang et al., 2018). From spiked-in DMSO control samples we know that the detection of higher gene numbers is possible with the developed pipeline. DMSO control samples were sequenced at a higher coverage (average of $7.7x10^6$ reads), since the number of droplets per experiment was higher. Accordingly, the number of detected genes (genes > 1 read) was on average 12216. The number of detected genes at the given coverage is comparable to the performance of previously reported high-throughput RNA-Seq based drug screens in plates (Ye et al., 2018). Taken together, an increased sequencing depth per drug combinations would allow the detection of higher gene numbers and subsequently give a higher sensitivity. A possible consequence is an improved detection of effects induced by drugs encoded by BC-bio. This is likely to improve correlations between drug signatures independent on the barcoding order, since we would expect the data across different screens to be less noisy. For this purpose we recently run a control experiment using only 4x4 combinations, resulting in 16 combinations out of which 10 are unique treatment conditions and the remaining 6 samples were

repetitions encoded with a different barcode combination **(Tab. 5.3)**. To additionally control for biases introduced by injection positions, we repeated the experiment with a changed order of barcodes, meaning BC-dT barcodes were used for the autosampler and BC-bio barcodes for the valve-module. From the results we expect to draw conclusions on whether an increased coverage improves the detection of effects from both drugs of combinations and whether we are able to exclude biases in the observed effects based on whether drugs are injected from the valve-module or autosampler.

Tab. 5.3: Combination control screen

	Imatinib	**Trametinib**	**YM115**	**DMSO**
Imatinib	Imatinib Imatinib	Imatinib Trametinib	Imatinib YM115	Imatinib DMSO
Trametinib	Trametinib Imatinib	Trametinib Trametinib	Trametinib YM115	Trametinib DMSO
YM115	YM115 Imatinib	YM115 Trametinib	YM115 YM115	YM115 DMSO
DMSO	DMSO Imatinib	DMSO Trametinib	DMSO YM115	DMSO DMSO

D I S C U S S I O N

6. General discussion and perspective

Finding the right combination is a challenging task when the number of possible choices is tremendous. That is clearly the case when considering all possible drug like molecules ($1x10^{30}$) present in the chemical space (Macarron et al., 2011), but yet a challenging task for all 1500 drugs approved by the FDA (Sun et al., 2013). Nevertheless, the financial (stating from clinical phase 2) and biomedical (higher efficacies, reduced toxicity and fewer side effects) advantages of defining new treatment strategies based on combinations of approved drugs, gives their discovery great potential. Due to the high numbers of possible combinations, the development of new approaches enabling the discovery of potential drug combinations is of great relevance. Massive sample reduction and increased storage capacity combined with a higher throughput, could potentially allow unbiased screens of drug combinations. Droplet-based microfluidics fulfils many criteria to push the boundaries for combinatorial drug screens. We believe that innovative new methods for combinatorial drug screens are of great potential for cancer research due to the emerging hallmark of inter and intra tumor heterogeneity and its impact on treatment strategies. The promise of combinatorial targeted therapies to prevent resistance has already led to several clinical trials and approvals for drug combinations (Al-Lazikani et al., 2012). However, limitations in defining new potent drug combinations do not only arise from the high number of potential combinations, but also the intertumour heterogeneity. The heterogeneity observed between tumors from different patient is mainly caused by genetic mutations within the tumor and can cause different susceptibility towards therapies (Vogelstein et al., 2013). Studies testing drug sensitivities directly on models derived from biopsies of solid tumors, such as patient derived cell lines (Crystal et al., 2014) or patient derived tumor xenograft mouse models (Gao et al., 2015) observed heterogeneous drug responses. Additionally, studies performing micro-titer based screens directly on blood cancer samples were performed to determine

patient tailored treatments (Pemovska et al., 2013) and to decipher genetics underlying drug responses (Dietrich et al., 2018). Patient derived models present a powerful tool to study drug responses on patient material, however, to what extent these models represent the full intratumour heterogeneity is difficult to assess. Pre-existing subclones responsible for tumor relapse due to drug resistance were described to be present at very low frequencies in pre-treated tumors (Kim et al., 2018) and therefore can be missed in genetic comparisons between primary tumors and their derived models. In order to enable drug screenings directly on patient material from solid as well as blood tumors a reduction in the required number of cells per screen could be beneficial and was shown to be achievable by the use of droplet based microfluidics (Eduati et al., 2018). Besides the possibility of studying interpatient heterogeneity, droplet-based microfluidics has the potential to enable better characterization of intratumour heterogeneity by screening drugs and drug combinations at the single cell level. The encapsulation of single cell into droplet along with drug combinations can be exploited to select individual cells based on their phenotype (e.g. resistance) followed by their characterization based on gene expression profiling. Implementation such as direct single cell sequencing on treated cells in droplets has the potential to give deep insights into the impact of heterogeneity on sensitivity towards large sets of drug combinations. In order to pave the way for such types of experiments, we developed novel microfluidic workflows for the generation of hundreds of combinations in picolitre-sized droplets, enabling screens on low input materials. In the presented work, up to 420 treatment conditions were tested on tumor cells encapsulated into droplets on the single cell level. Additionally, we provide the first framework for highly multiplexed gene expression readout of cells perturbed by drug combinations using drug specific barcoding of transcriptomes.

6.1. Indexing of droplets with barcode combinations for multiplexed droplet experiments

In chapter 3 we presented a novel approach to generate combinations in picolitre-sized droplets by using a valve based microfluidic pipeline. The control of injected reagents by valves allowed rapid switching between 14 sample inlets and their mixing. Combinations of two reagents were encapsulated into droplets on a

separate drop-maker chip and a valve based sample collector controlled their collection. We implemented a novel DNA-based barcoding approach that enables the identification of each combination and therefore all droplets could be stored in the same tube. The generation and the DNA-barcoding of combinations in droplets overcame the limitation of other valve-based approaches to store combinations sequentially in large droplets, which is not fully scalable (Eduati et al., 2018; Rane et al., 2015). Furthermore, it enables the integration of single-cell phenotypic screens, e.g. to isolate and characterize resistant cells in a multiplexed fashion. In contrast to previous approaches, in the presented work, combinations were generated in picolitre-sized droplets and stored in a single tube. As compared to fluorescence-based barcoding used for encoding single compounds (Brouzes et al., 2009) or drug concentration (Miller et al., 2012), the presented DNA-based barcoding approach can be scaled up to encode millions of conditions. The generation of combinations from the valve-module allowed production of 49 combinations with 2100 droplets per condition in around 12 min and thus presents a fast way to test several conditions in droplets. Since each condition can be demultiplexed by its barcode combinations, all droplets with diverse contents can be processed in a single emulsion. Multiplexed screens increase the throughput compared to conventional droplet-based screening approaches by processing all conditions in a single run. Furthermore, the possibility to multiplex several conditions has the advantage that less input material is required since only several thousand droplets per conditions can be processed. As an example we performed one droplet sorting experiment with a pool of 49 conditions (only different DNA barcodes) instead of sorting for individual 49 conditions, providing a method to increase throughput of droplet based sorting screens. A subset of droplets with barcode combinations was additionally labelled with a fluorescence dye. Droplets of interest were sorted based on their fluorescence signal and the enrichment was measured by comparing sequenced barcodes obtained from sorted droplets and unsorted droplets. In future, this setup can be applied to screen for potent combinations of perturbants, such as anti-cancer drugs. Sorting based on reporter signals would allow the enrichments of combinations showing the desired effects. By exposing cancer cells to drug combinations and subsequent sorting of droplets based on fluorescence apoptosis readouts, potent combinations against cancer can be identified. Performing such experiments on droplets generated from the

workflow described in chapter 4, would allow to separate heterogeneous cell populations into responding and non-responding populations

As a strategy to increase the chemical diversity in droplets, compounds injected into the valve-module could be diluted. One possibility is to generate multiple discrete concentrations of drug combinations by diluting each sample plug with a defined volume injected from a third valve. Additionally, by using different valve opening times, the ratio in which two reagents are combined can be altered. In a more advanced approach, the increased diffusion due differences of flow velocity in a tube (Taylor-Aris dispersion) could be exploited to generate dilutions of compound plugs as described by Miller et al. Here, the use of a miscible carrier phase was used to dilute compounds, resulting in bell-shaped concentration profiles. More complex dilution patterns could be generated by opening two valves in different sequential orders and the use of an additional valve for dilution. The generation of continuous dilution curves could help to increase the sample space and could serve as a powerful tool to find synergistic drug pairs. However, more challenging to implement is a readout for each concentration using DNA-barcoding approach. A possible solution could be the ligation of three barcodes where two are used to encode a combination and one is used to encode a concentration. However, this would need one barcode for each concentration, which would again be limited by the number of available valves. A second approach could be to implement fluorescence dye based readout to encode dilutions of compounds as shown by Miller and colleagues (Miller et al., 2012). We could moreover increase the number of combinations by the generation of all possible pairs (91 pairwise combinations from 14 valves) from the valve-module. For this purpose each compound could be encoded by one barcode sequence present as a mixture of two barcode species so that all barcodes become compatible with each other and such that all valves can be combined with each other. A pipeline that additionally provides the option to generate all possible combinations out of three compounds would generate up to 364 combinations without increasing the number of sample inlets. Along with implementations of the discussed improvements, the pipeline will exceed multiplexing capacities for droplets in emulsions of previously presented work. Therefore, the valve-based approach together with DNA based barcoding of samples, presents a powerful technology for screening drug combinations using droplet-based microfluidics.

A long lasting limitation for drug screens in emulsion is the exchange of drugs based on micellular transport (Gruner et al., 2016). In order to exploit the full potential of droplet-based microfluidics for combinatorial drug screen, a crucial step is to develop technologies preventing exchange of compounds between droplets. A potential solution, we tried to integrate into the pipeline, is the use of droplet arrays in which droplets are stored individually in traps. This allows the removal of surfactant since droplets are spaced out and thus do not coalesce. The advantage of this approach was exploited in a recently published work to perform large-scale combinatorial drug screens in droplets (Kulesa et al., 2018). However, it requires high number of microfluidic devices for droplet storage and processing of droplets after incubation is challenging since droplets stick to traps in the droplet array. We also tested the so-called pickering emulsions to prevent surfactant mediated exchange between droplets. Here, fluorinated amphiphilic nanoparticles are used to stabilize the water-oil interface (Pan et al., 2015). By use of fluorinated nanoparticles the formation of micelles and consequently exchange can be prevented. Until now, we did not manage to synthesize nanoparticles that stabilize water in oil droplets over a long time period. Since long incubation times can result in an exchange-mediated equilibrium of drug concentrations between droplets (Gruner et al., 2016), we focused on drug screens using RNA-Seq readouts. Here, as compared to cell viability readouts, shorter incubations are sufficient (6h) to determine the effects of drugs on gene expression (Subramanian et al., 2017). Another advantage in this context is that gene expression readouts allow detection of drug exchange by comparing expression patterns between different treatment conditions, especially of untreated controls.

6.2. Gene expression based profiling of drug combinations using droplet-based microfluidics

The following section discusses the results and conclusions drawn from chapter 4 and chapter 5. Here we introduced a microfluidic pipeline to generate drug combinations from an extended version of the valve-module and an autosampler. As compared to an approach previously described by our group (Eduati et al., 2018), we could significantly increase the number of possible combinations (420 instead of 55). Moreover, in the presented work all 420 combinations were stored

in a single tube, which reduces screening costs and increases the storage capacity. Another advantage is the possibility to process droplets after incubation by standard methods used for droplet-based microfluidics, such as droplet sorting or pico-injection, which allows selection of droplets or addition of reagents, respectively. This is particularly interesting considering that the use of picolitre-sized droplets allows the encapsulation of single cells along with drug combinations and thus processing of droplets allows single cell phenotypic readouts (e.g. apoptosis). Pooling of all compound combinations in a single tube was again facilitated by DNA-based barcoding of each combination. In this case, we used barcode combinations additionally functionalized with poly-dT sequences facilitating barcoding of transcriptomes according to the combination of drugs cells were exposed to. Sequencing of barcoded transcriptomes was applied to assign gene expression data to each treatment condition. The barcoding approach is deterministic, meaning each barcode identifies a specific treatment condition. This is in contrast to random barcoding as used for droplet-based single cell sequencing approaches (Klein and Macosko, 2017). Therefore, the presented work has the potential to perform highly multiplexed gene expression experiments from 420 treatment conditions using a semi-automated workflow. To the best of our knowledge, the presented approach is the first one allowing the on-demand generation of such high numbers of combinations, which can be stored in a single emulsion of picolitre-sized droplets. Additionally, the integration of a RNA-Seq compatible barcoding strategy with the developed droplet-based screening pipeline is the first approach that aims at a highly multiplexed gene expression based readouts of drug combinations. This is in contrast to a recently published work by Kulesa and colleagues (Kulesa et al., 2018), where drug combinations were generated in stationary droplet-arrays by fusion of two droplets. All droplets were stored in the droplet-arrays over the time of incubation after which growth rates of bacteria were measured by a single fluorescence raedout. In order to screen 100,800 drug combinations 126 microarray chips were used with only 13 droplets per combinations.

The translation from micro-titer plate screens towards emulsion-based screens harbours great potential for several future applications. The relative low numbers of cells necessary for 420 combinations with maximal 1000 cells per sample may be exploited to perform functional screens directly on patient material. Due to low

incubation time necessary to obtain gene expression profiles from drug treatments, no long term culturing would be required. Gene expression profiles from the LINC database were prepared from cells exposed to drugs for 6h (Subramanian et al., 2017). Additionally, we detected drug-specific clustering of gene expression data prepared from cells exposed to drugs for only 6h (data shown in appendix Fig. 9.2) and 12h (Fig. 5.3). Consequently, incubation times of only 12h or even 6h may be sufficient to draw conclusions about drug sensitivity based on gene expression data. Fast prediction on drug sensitivity of primary tumor cells can be beneficial to prevent selection of specific subclones and to preserve their phenotypes in terms of drug sensitivity (Montero et al., 2015). The possibility to use gene expression under perturbation to predict anti-cancer drug sensitivity was recently demonstrated by the integration of perturbation signatures from the LINCS database and cell viability screens (Szalai et al., 2018). Szalai et al. defined consensus cell death signatures, which were used to predict cell viability. This may turn out as a useful method to additionally predict synergistic drug pairs based on cell death signatures scores exceeding the additive scores of the two compounds alone. Taken together, gene expression data obtained from cells after short incubation times can be used to determine potent drug combinations, which eventually will improve drug screens, in particular for primary tumor cells. Moreover, the gene expression data of perturbed cells is a more informative readout of drug sensitivity as compared to reporter-based readouts, such as cell viability or apoptosis alone. Besides the possibility to determine cell death signatures of drug combinations, the gene expression data can be used to define underlying signalling pathway activities (Schubert et al., 2018). Knowledge on the pathway activity of perturbed cells can be informative to make associations between synergistic actions of drug pairs and pathway activity.

In the presented work, cell suspensions were used at densities suitable for obtaining single cells in droplets. The generation of droplets harbouring single cells provides the possibility to introduce droplet-processing methods for functional single cell characterization. We believe that this is of great potential for studying the impact of tumor heterogeneity on drug sensitivity, since effects of individual subpopulations can be studied. Quantifying the number of non-responding cells towards drug combinations could allow prediction on resistance subpopulations. Such screens can be performed by encapsulating cells with fluorescence reporter

assays for apoptosis (e.g. Caspase 3/7 substrates) followed by sorting of droplet containing viable cells (i.e. apoptosis negative). Enrichment of surviving cells and their characterization by gene expression can provide interesting insights on non-responding or drug resistance phenotypes in a highly multiplexed fashion. We have demonstrated in section 3.7 the possibility to enrich a specific droplet population and determine their content by sequencing the enriched barcode combinations. In a next step, sorting of droplets can be combined with the later drug screening to enrich for non-responding cells. However, to fully exploit the use of the presented pipeline for screens on biopsies from solid tumors and to sort those cells based on apoptosis, we would need to implement strategies to prevent programmed cell death induced by the loss of adhesion (Frisch and Screaton, 2001). To overcome this limitation, a possibility could be cell encapsulation into droplets along with polymers used in 3-D cell culture techniques as previously reported (Sart et al., 2017; Yu et al., 2010).

Furthermore, the encapsulation of single cells along with drug combinations and their corresponding barcodes can be exploited in single cell sequencing methods. Our current pipeline would in principle allow the co-encapsulation of barcoded beads and cells as described in the Drop-Seq protocol (Macosko et al., 2015). Instead of lysing the cells directly, we could apply pico-injection as described in section 2.7 to lyse cells after their incubation together with drug combinations. Hence, we could sequence the transcriptome of each cell individually and obtain single cell resolution of responses to drug combinations. To additionally barcode each cells transcriptome according to the drug combination it was exposed to in the droplet, we would need to adapt our combinatorial barcoding approach. One possibility to encode the drug information together with single cell transcriptomes, is to apply hybridization of barcodes harbouring poly-dA sequences to barcoded beads as it is already described by for single cell multi-omics methods (Stoeckius et al., 2017).

Using single cell sequencing together with the introduced microfluidic pipeline would additionally allow the implementation of a new barcoding strategy for encoding drug concentrations. We could dilute compound plugs by opening an additional valve to inject a DMSO control at a certain ratio. However, determining the concentrations by sequencing RNA-Seq is challenging since we would only determine the number of reads with a particular barcode combination. Introducing

an additional barcode for each droplet (i.e. cells), as it is done in Drop-Seq, could potentially be exploited to count the numbers of drug barcodes assigned to each cell barcode. Extending drug barcodes with unique molecular identifiers (Islam et al., 2014) would allow the quantification of drug barcodes assigned to a cell barcode. As a result, concentration of drug barcodes could potentially be measured by sequencing and consequently the concentration of drugs and its influence on gene expression could be assessed. Diluting compound plugs from the valve-module would increase the overall number of possible combinations without expanding the current setup. Screen over five different distinct concentrations would already result in 2100 treatment conditions using the current drug panel. Performing single cell sequencing from hundreds of drug combinations will be a powerful tool to elucidate their effects on gene expression. We envision that implementations of such technologies will provide information of why subpopulations do not respond to drug treatments and allow definition of drug combination that overcome resistance mechanisms.

M A T E R I A L & M E T H O D S

7. Material and Methods

7.1. Methods in Microfluidics

7.1.1. Mask design

Photo-masks are necessary in the process of standard photolithography. The design of a photo-mask was drawn in AutoCAD (Autodesk Inc., USA). Designs were printed (Selba S.A. Versoix Switzerland) on transparent plastic slides at 25400 dpi. Depending on the later application designs were either printed as positives or negatives. Positives were used for mould manufacturing from positive photoresist for the production of microfluidic devices used in the valve-module. Negative prints of designs were used for producing moulds from negative photoresists (see below).

7.1.2. Photolithography for mould manufacturing

Negative Photoresists

Silicon-wafers (Siltronix, Silicon Materials, Germany, 3" or 4") were first cleaned using an air gun and then placed on a hot plate at 140 ˚C for 10 min. In the next step wafers were coated with a layer of SU-8 2075 photoresist (MicroChem Corp., Newton MA): A wafer was placed on a spin coater (Laurell Technologies Corp., North Wales PA), photoresist was poured on the wafer and evenly distributed by spinning the wafer at a given speed. The time and speed of spinning determined the thickness and thereby the channel height of the mould and these parameters were adapted according to the manufacturer's instructions. Coated wafers were baked at 65 ˚C and 95 ˚C for times given by the manufacturer for a given photoresist thickness. A negative of a photo-mask was placed on the wafer and exposed to UV-light in a mask aligner (Karl Suss MA45). Wafers were again baked at 65 ˚C and 95

˚C followed by the development with mr-Dev-600 (micro resist technology GmbH, Berlin) in order to remove all areas that were not exposed to UV-light. After a last baking step at 150 ˚C the moulds were ready for microfluidic chip manufacturing (see below).

Positive Photoresist

Positive photoresists were used to obtain rounded channels used for chips in the valve-module. Silicon-wafers were cleaned and baked at 140 ˚C for 10 min. AZ 40XT photoresist (MicroChemicals GmbH, Ulm) was poured on a wafer, which was distributed in a spin coater at 1300 rpm for 40 sec. After a soft-baking step **(Tab. 7.1)** a positive photo-mask was placed on top of photoresist layer, which was exposed to UV-light for 150 sec. Wafers were left at RT for 5 min followed by a post-exposure baking step **(see Tab. 7.1)**. Moulds were developed using AZ 726 MIF Developer (MicroChemical GmbH, Ulm) to remove all UV-exposed areas from the wafer. In order to obtain rounded channels, wafers were baked according to the round bake step **(see Tab. 7.1)**. Subsequently moulds were ready to produce microfluidic chips used for valve-modules (see below).

Tab. 7.1: Baking times for positive photoresist

Time [min]	4	2	2	1	2	2	2	4
Soft bake	55 °C	65 °C	75 °C	85 °C	95 °C	105 °C	115 °C	
Post-exposure bake	55 °C	65 °C	75 °C	85 °C	95 °C			
Round bake	55 °C	65 °C	75 °C	85 °C	95 °C	105 °C	115 °C	125 °C

7.1.3. PDMS Membranes

Polydimethylsiloxane (PDMS) and curing agent (Sylgard 184 silicone elastomer kit, Dow Corning Corp., USA) were mixed at a ratio of 10:1 w/w and placed in a desiccator until all air bubbles were removed. The mix was distributed on transparency slides using a spin coater at 500 rpm for 19 sec. Membranes were cured over night at 65 ˚C.

7.1.4. Microfluidic chip manufacturing

Moulds were placed in petri dishes and filled with a 10:1 mix of PDMS and curing agent. After desiccation, moulds were baked over night at 65 ˚C. On the next day the PDMS was cut and peeled off the mould. Inlets to insert tubing or electrodes were

punched using biopsy punches with a diameter of 0.75 mm. The PDMS was cleaned using air pressure and adhesive tape and then plasma bonded (Femto, Diener electronic GmbH + Co. KG, Germany) either to glass slides (negative photoresist) or PDMS membranes (for all chips with valves). Subsequently, devices were kept at 65 ˚C for 5 min. Chips for the valve-module were additionally bound to a glass slide only covering the inlet and outlet ports. This was necessary to prevent rupture of the membrane and thus leakage when inserting tubings into the ports. In order to obtain hydrophobic channel surfaces, Aquapel (Autoserv, Germany) was flushed through the channels and subsequently removed using an air filled syringe.

7.1.5. Preparation of microfluidics chips with horizontal inlets

Valve-module chips and drop-makers used for producing high numbers of combinations, were connected by a delay tubing which was inserted from the side allowing reagent injection parallel to the channels. This minimizes cross-contamination due to plug breakage (Clausell-Tormos et al., 2010). Holes for valve-module outlets were punched from the side using a 0.5 mm biopsy punch. After plasma-bounding the chip to a PDMS membrane, a piece of Polytetrafluoroethylene (PTFE) tubing (UT3, Adtech Polymer Engineering LtD, UK) with an outer diameter of 0.4 mm was inserted into the side inlet until it reached the funnel structure of the outlet channel. Chips were bound to glass slides big enough to cover parts with inlets and outlets, preventing membrane rupture by the insertion of tubing. The tubing inserted into the outlet channel was bound to the chip using UV curing adhesive (Loctite, Henkel AG & Co. KGaA, Germany). Drop-maker chips were produced by first punching holes with a 0.75 mm biopsy punch for oil and cell inlets and the outlet. Insertion ports for the autosampler were punched using a 1.25 mm punch and inlets for the delay tubing using 0.5 mm punch Holes for inserting the delay tubing were punched from the side to prevent plug break-up. The drop-maker devices were first bonded to a PDMS membrane and then to a glass slide. After aquapel treatment chips were ready to use.

7.1.6. Injection of fluids into microfluidic devices

The injection of fluids into channels of a microfluidic device was performed using syringes mounted on syringe pumps. PTFE tubings with in outer diameter of 0.76 mm (HW30, Adtech Polymer Engineering LtD, UK) were connected with Luer-Lok

syringes (Becton Dickinson, USA) using 27G ¾ gauge needles (Becton Dickinson, USA). Syringes were filled with reagents and mounted on PHD 22/2000 Syringe Pumps (Harvard Apparatus, USA). In the pumps' menu the syringe diameter and flow rates were set accordingly. The free end of the tubing was connected with a microfluidic device via the punched tubing inlet and thereby reagents were injected at a defined flow rate into the channels of the device.

7.1.7. Braille display operations

A SC-9 braille display with 64 pins (KGS Corporation, Japan) was connected over communication port to a computer. The sample on demand software (Eduati *et al.*) was used to control the pins of the braille display. All pins were automatically moved down once the braille display was connected. For using the braille display to direct injected fluids into different channels of the microfluidic device, always two pins were operated at the same time. In the default mode the outer pins were moved upwards and the two lines of pins in the middle were moved downwards **(Fig. 7.1A)**. In this mode, all liquids injected into the chip are directed towards the waste outlet since pins blocked the channels directing fluids the outlet for plug production. Injected fluids were directed to the outlet for plug production by closing the waste valve (pin in the middle row) and by opening the corresponding pin in the outer row **(Fig. 7.1A)**. This process was automated by uploading a CVS file in the sample on demand software, which contained a sequence of valve operations and opening times.

7.1.8. Setting up a valve-module

The channels of a valve-module chip (16 valves or 24 valves) were filled with trypan blue (Thermo Fisher Scientific Inc., USA) using a syringe. The braille display was connected to a computer and the valve structures of the chip were aligned on top of the pins **(Fig. 7.1B)**. A plexiglass holder was used to fix the chip on the pins. Actuating one valve in the sample on demand software started the priming mode in which all waste valves were open and sample injection valves were closed. Blocking channels was achieved by pins pushing the elastic membrane against the channel walls. First pieces of tubing were inserted into the waste outlets.

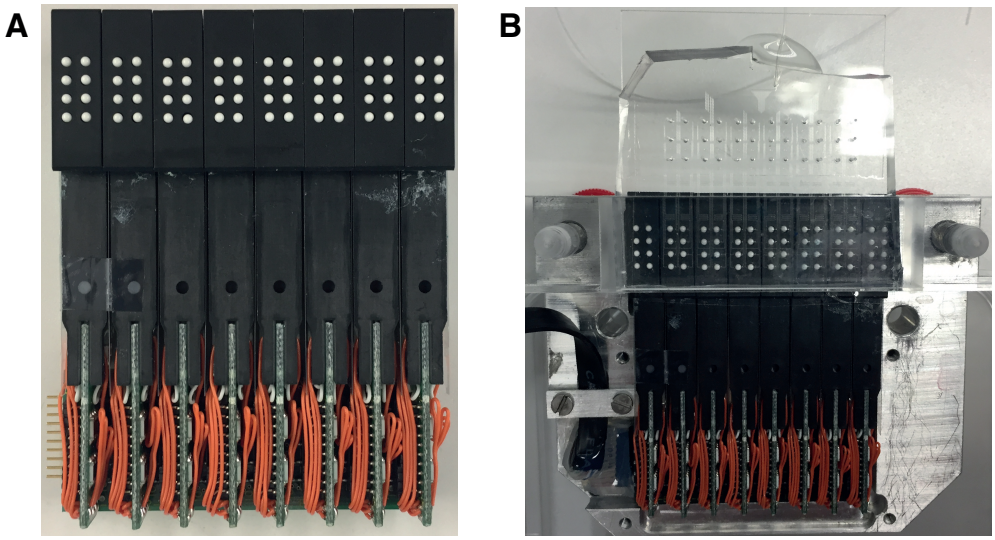

Fig. 7.1 Pictures braille display and valve-module

(A) Picture of a SC-9 braille display used in this work. The actuation of the white pins was controlled by a LabVIEW software (sample on demand). **(B)** Picture of a valve-module: A PDMS chip was aligned on top of the white pins of a SC-9 braille display. Chip was fixed by a plastic bar. Moving the pins up and down allowed closing or opening of the channels above.

Subsequently, the tubings connected to 5 ml syringes filled with aqueous solutions were inserted into the inlets of the valve-module device. Additionally, 5 ml syringes filled with Novec™ 7500 oil (3M Company, St. Paul MN) were connected with inlets of the valve-module chip. All syringes were mounted on syringe pumps and fluids were injected at 500 µl/h.

7.1.9. Setting up a sample collector module

A sample collector chip bond to a PDMS membrane was filled with trypan blue. The chip was placed on a Braille display with each channel being on top of a row of four pins. The Braille display was connected to an in-house LabVIEW (National Instruments, USA) software that allowed the simultaneous actuation of all four pins in one row. The two rows below the two channels were selected and one was defined as waste channel and one was defined as collection channel. Closing of one channel was synchronized with the opening of the other channel. Fluorescence threshold and time windows were set and determined when and for how long the collection channel was opened. The two pieces of tubing from the drop-makers' outlets were connected with two inlets on the sample collector. Two short pieces of

tubing were inserted into the outlets of the sample collector to direct droplets to a collection and waste tube.

7.1.10. Autosampler operations

The autosampler (Dionex UltiMate 3000 Analytical Autosampler WPS-3000 SL, Thermo Fischer Scientific Inc., USA) was operated using Chromeleon software (Thermo Fisher Scientific Inc., USA). In an operation sheet the wells of a 96-well plate were selected and compounds were aspirated by the autosampler accordingly. PBS was used as a carrier fluid and was injected by an external syringe pump into the autosampler valve at 500 µl/h and was used to inject aspirated compounds into the drop-maker. The outlet of the autosampler was connected to a PEEK tubing (0.12 mm inner diameter, Agilent Technologies Inc., USA) that was inserted into the autosampler inlet of a drop-maker. After each compound the fluidic system of the autosampler was rinsed by an internal syringe pump with 400 µl PBS. For each experiment the waiting time between each injection was adapted in the command sheet of Chromeleon software to match the production and injection times of compound plugs from a valve-module.

7.1.11. Optical set-up used for measuring fluorescence intensities of plugs and droplets

Spectroscopic fluorescence measurements were performed using an in-house optical system. Laser beams with wavelengths of 375 nm, 488 nm and 561 nm were directed over dichroic mirrors into the objectives of an inverted light microscope (Eclipse Ti-S, Nikon GmbH, Germany). For all measurements a 40x objective was used. The emitted light was directed through bandpass filters on a set of three photomultipliers (PMTs) and allowed simultaneous measurements of three fluorescence dyes.

7.1.12. Cell encapsulation into droplets

K562 cells were cultured in IMDM media (Thermo Fisher Scientific Inc., USA) supplemented with 10 % fetal bovine serum (FBS, Thermo Fisher Scientific Inc., USA) and 1 % Pencillin Streptomycin (Thermo Fisher Scientific Inc., USA). For cell encapsulation experiments, K562 cells were washed in PBS and resuspended in Gibco FreeStyle™ 293 Expression Media (FS media, Thermo Fisher Scientific Inc.,

USA) with 4% FBS at a concentration of $2x10^6$ cells/ml. K562 cell suspensions were loaded into a 3 ml syringe together with a magnetic mixing disc (VP Scientific, USA). Syringe was mounted on a syringe pump with the needle facing downwards and connected with the cell inlet of a drop-maker. To prevent cell sedimentation, the magnetic disc was rotated at a low speed using a magnetic stirring system (VP Scientific, USA). In all cell-based experiments done for this work, injected cells were diluted on the chip by other aqueous phases down to a concentration of $5x10^5$ cells/ml.

7.1.13. Microfluidic pipeline for combining compounds injected into the valve-module

This chapter aims at describing the general set up of the microfluidic pipeline described in chapter 3. A valve-module with 16 inlets was set-up and connected with 14 syringes containing barcodes and fluorescent dyes and 2 syringes filled with Novec oil. All liquids were injected at 500 µl/h into the valve-module and valves were operated using a CVS file with the following sequence for combining valves A to G with valves 1 to 7:

Valves	Opening times	Plugs
A + 1	5s	Collection Plug
2 x oil	2s	Oil Spacer
A + 2	3s	Waste / Washing Plug
2x oil	5s	Oil Spacer
A + 2	5s	Collection Plug
…	…	…
G + 7	5s	Collection Plug

A short tubing was used to connect the outlet of the valve-module chip with the braille inlet of the drop-maker, which was placed under the microscope. The second inlet of the drop-maker chip was used to continuously inject the 3x ligation mix at 500 µl/h. QX200™ droplet generation oil (BioRad Laboratories Inc., USA) was injected at 4500 µl/h to produce droplets with a volume of approx. 800 pl. The sample collector was connected over two short pieces of tubing with the outlets of the drop-maker. Fluorescence intensities of incoming plugs were measured

upstream of the flow-focussing junction and were used to control droplet collection. For a 5s collection plug, a delay time of 500 ms was followed by a collection time of 4 sec. The waiting time was set to 8.6 sec to ensure that the system did not collect any droplets generated from washing plugs, but was ready for the next cycle before the subsequent collection plug was injected. One cycle of 49 combinations with 5s opening times took 12 min and 15 sec and produced approx. 2000 droplets per combination. For experiments with varying collection times, the time for droplet collection was adjusted accordingly.

7.1.14. Microfluidic pipeline for generating combinations out of compounds from the valve-module and autosampler

This chapter aims at describing the general set-up of the microfluidic pipeline used in chapter 4. Here, a valve-module with 24 inlets was used to generate combinations with compounds from a 96-well plate. 20 5 ml syringes filled with compound-barcode mixtures stained with fluorescence dyes were injected into the valve-module (valves 3 to 22). The two outermost valves on both sides were used to inject Novec oil supplemented with 0.1% 1H,1H,2H,2H-Perfluorooctanol (PFO; abcr GmbH, Germany) to prevent wetting at the channel walls. All liquids were injected at 500 µl/h and compound plugs were produced by the valve-module using a CSV file for the sample on demand software with the following valve opening sequence:

Valves	Opening times	Plugs
A	7 sec	Compound Plug
4 x oil	2 sec	Oil Spacer
B	7 sec	Compound Plug
4 x oil	2s	Oil Spacer
…	…	…
T	7 sec	Compound Plug
2 x oil	~120 sec	Plug injection

The opening times for compound plugs can be varied in order to produce more droplets from each compound combination. The opening time for plug injection was adapted for each experiment to match the time necessary to inject all

compound plugs. The length of the delay tubing (connected to the valve-modules outlet) was determined, by injecting all 20 compound plugs into the delay tubing. Then the tubing was cut before the point were the first compound plug was located and inserted from the side into braille inlet of the drop-maker. The drop-maker chip was placed under the microscope to measure fluorescence emissions. The cell inlet of the drop-maker was connected with a 3 ml syringe filled with a single cell suspension at $2x10^6$ cells/ml, which was injected at 500 µl/h. The autosampler inlet was connected with the autosampler and carrier fluid was injected at 500 µl/h. Novec oil with 1% Pico-Surf1 surfactant (Sphere Fluidics, UK) connected with the oil inlet was injected at 6000 µl/h in order to produce droplets of approx. 800 pl, which were collected in a 15 ml tube (Becton Dickinson, USA) on ice. Once the set-up was running, the sample on demand software was set to generate 21 cycles of 20 compound plug and their injection at 1000 µl/h (2 x oil at 500 µl/h). Once five compound plugs (A-E) of the first cycle were produced the autosampler-based injection of compounds from a 96-well plate was started. Thereby, we ensured that the plateau concentration of compound injected from the 96 well-plate was reached once the first compound plug was injected into the drop-maker. Combinations of compounds from the braille display with 21 compounds from the autosampler were produced and collected in a single tube.

7.1.15. Picoinjection for cell lysis and barcode ligation

A picoinjection device was placed on a hotplate at 95 °C until the PDMS was heated up. Subsequently, low melting solder was inserted into the ports for the electrodes, which were connected with cables **(Fig. 7.2A)**. The channels of the device were treated with Aquapel. Electrodes were connected to the high voltage amplifier and ground. Novec supplemented with 1% surfactant (Pico-Surf 1, Sphere Fluidics) was injected into the oil inlet to flush all channels with oil. A droplet emulsion was loaded into a 3 ml syringe connected to a piece of tubing. Syringe was mounted on a syringe pump with the needle facing upwards and droplets were injected at 500 µl/h. Once all the air was removed from the tubing, it was connected to the droplet inlet and injected droplets were spaced out by the oil. At the pico-injector inlet a solution for cell lysis and ligation (Ligation and Lysis mix) was injected at 200 µl/h and cooled using ice bags. A continuous electrical field of 0.1V was applied using a function generator.

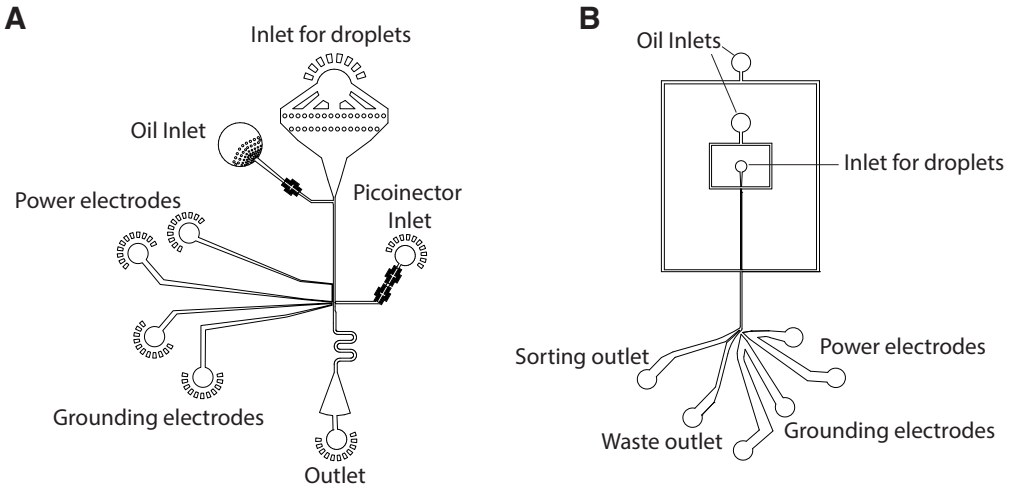

Fig. 7.2: Chip designs for Pico-injections and Sorting of droplets

(A) Chip design used to manufacture moulds for Pico-injection devices. Channels for electrodes were filled with solder and cables were inserted into each port. **(B)** Designs for manufacturing sorting devices. Electrodes were filled with solder and ports were connected to cables

This resulted in the injection of ligation and lysis reagents into each droplet passing by the nozzle of the pico-injector. Droplets were collected on ice in a 50 ml tube (Becton Dickinson, USA).

7.1.16. Fluorescence activated droplet sorting

A soring device was placed on a hotplate at 95 °C and once the PDMS was heated up, low melting solder was inserted into the electrode ports and connected with cables **(Fig. 7.2B)**. Two syringes filled with Novec were connected to the oil inlets and channels were flushed with oil. The chip electrodes were connected with the high voltage amplifier and ground. Droplets were loaded into an 1 ml syringe connected to a piece of tubing. The syringe was placed on a pump and droplets were injected at 100 µl/h and the flow rates of the two oil injections were adjusted to space out droplets. The waste outlet of the sorting device was connected with short piece of tubing (TW24, Adtech Polymer Engineering Ltd, UK) and the collection outlet to a slightly longer piece of tubing (HW30). The differences in resistance between the two connected tubings caused all droplets to enter the waste channel. We used an in-house LabVIEW software to measure peak fluorescence intensities of droplets and to set gates on the droplet population of interest. When a droplet of interest passed the laser, the software applied an

electric pulse that pulled the droplet into the collection channel based on positive dielectrophoresis.

7.1.17. Breaking emulsions

In order to extract droplet contents, it was necessary to break the emulsion by inducing the coalescence of all droplets. For this purpose, oil at the bottom of the tube was removed using a 1 ml pipette. Subsequently, 1 ml of PFO was added to an emulsion followed by a couple of firm shakes. When all droplets were broken the aqueous phase was pipetted into a fresh tube, otherwise more PFO was added.

7.2. Preparation of samples for the valve-module and autosampler

7.2.1. Fluorescence dyes

For the validation of the microfluidic pipelines used in this work, solutions with fluorescent dyes were used. Cascade Blue hydrazide trilithium salt (Thermo Fisher Scientific Inc., USA) was dissolved in DMSO and diluted in PBS to 20 µM and then loaded into 5 ml syringes for the injection into inlet of a valve-module. Fluorescein sodium salt (Sigma-Aldrich Corporation, USA) was dissolved in water and diluted to 10 µM in PBS. Fluorescein solution was aspirated into 5 ml syringes for the injection into inlet of a valve-module.

7.2.2. Barcode design

Barcodes encoding droplet contents

The 10 bp barcode sequences were designed by nxCode (Hannon Lab) and further processed by adding ligation sites to the 3'-end and PCR-handles to the 5'-end (**Tab. 7.2**). Fwd sequences of BC-1 to BC-7 were ordered with a biotinylation site at the 5'-end. Reverse complements (Rev) without ligation sites were designed using Wily DNA Editor (by A. Untergasser) and ordered with a 5'-end phosphorylation site to enable ligation. All sequences were purchased from Eurofins Genomics.

Tab. 7.2: Barcode sequences for droplet barcoding

BC-1 to BC-7	Fwd: [BIO]CTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNgcggc
	Rev: [Phos]NNNNNNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAG
BC-A to BC-G	Fwd: CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNNNNgccgc
	Rev: [Phos]NNNNNNNNNNAGATCGGAAGAGCACACGTCTGAACTCCAG

Barcodes for combinatorial barcoding with RNA-Seq

The 10 bp barcode sequences were generated using bgen (gear.embl.de), a tool that generates random DNA-sequences with balanced base distribution and defined length. The sequences of BC-bio Fwd were attached to a primer handle at the 5'-end and a ligation site at the 3'-end (**Tab. 7.3**). Reverse complement strands (Rev) without ligation site were designed using Wily DNA Editor (by A. Untergasser). BC-bio rev sequences were phosphorylated at the 5'-end. Fwd strands of BC-dT barcodes had a poly-desoxythymidin (poly-dT) sequence at the 3'-end and were phosphorylated at the 5'-end. The BC-dT rev sequences were attached to a ligation site complementary to the one of BC-bio Fwd sequences. All sequences were purchased from Eurofins Genomics.

Tab. 7.3: Sequences for RNA-Seq based combinatorial barcoding

BC-bio	Fwd:	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACNNNNNNNNNNgcggc
1 to 22	Rev:	[Phos]NNNNNNNNNNGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-dT	Fwd:	[Phos]NNNNNNNNNNTTTTTTTTTTTTTTTTTTTTVN
A to T	Rev:	NNNNNNNNNNgccgc

7.2.3. Barcode annealing

Sequences were dissolved in nuclease-free water (Thermo Fisher Scientific Inc., USA) at 100 µM and stored at -80 °C. Complementary sequences (Fwd + Rev) were mixed at equimolar concentrations and diluted to 20 µM using nuclease-free water. Mixtures were placed in a thermal block at 95 °C for 10 min and subsequently kept at room temperature for 1 h. 20 µM stock were stored at -20 °C.

7.2.4. Samples for the generation of barcode combinations

All annealed barcodes (14) were diluted in water to 75 nM. BC-A to BC-G were supplemented with 20 µM Cascade Blue for plug detection by the sample collector. Barcode solutions were loaded into 5 ml syringes connected to tubing for the injection into a valve-module. For fluorescence activated sorting of barcoded

droplets, BC-3 solutions were supplemented with 40 µg/ml of Alexa Fluor 488 (Thermo Fisher Scientific Inc., USA).

7.2.5. Barcode drug samples used for valve-module based injection

All compounds were dissolved in DMSO and stored at -80 °C. Drugs used for valve-module based injections were pre-diluted to 40x of their GR20-concentration measured for K562 cells. Working solutions were prepared by diluting drugs 1:20 in FS media and by adding BC-dT at 1 µM to their assigned drugs. Drug barcode solutions were stained with Cascade Blue or Fluorescein, loaded into 5 ml syringes and injected into the valve-module.

7.2.6. Barcode-drug samples for the autosampler-based injection

Drugs that were aspirated from a 96-well plate for the autosampler-based injection were pre-diluted in 96-well plates at 80x of their GR20-concentration measured for K562 cells. Drugs were diluted in 96-well plates to 4x the GR20-concentration using FS-media and BC-bio solutions. Barcodes were added to their assigned drugs at 4 µM and each drug barcode mixture was labelled with 10 µM Fluorescein.

7.3. Sequencing library preparations

7.3.1. Prepare barcode libraries for sequencing

For ligation of barcodes, droplets containing a set of two barcodes were incubated at room temperature for 30 min. Afterwards the emulsion was broken with PFO and the aqueous phase was transferred into a fresh tube. C1 dynabeads (Thermo Fisher Scientific Inc., USA) at 5 µg/µl in 6x SCC buffer (Thermo Fisher Scientific Inc., USA) were added at a ratio of 1:1 (v/v) to the ligated barcodes. For the binding of barcodes to the beads, the tube was incubated for 20 min under rotation at room temperature. Beads were placed in a DynaMag™ magnet (Thermo Fisher Scientific Inc., USA) and were washed once in TE-SDS and two times in water and subsequently re-suspended in water at 10 µg/µl. Barcodes were amplified using Phusion PCR mix and indexed primer sequences (NEXTFlex 16s V1-V3 Amplicon-Seq Kit, Bioo Scientific Corp., USA) . PCR products were purified on 1.8 x the volume of Ampure XP beads (Beckman Coulter, USA) according to the manufacturers

instructions. Libraries were pooled and barcodes were sequenced in one direction on a MiSeq system (Illumina Inc., USA) together with PhiX spike-ins (Illumina Inc., USA).

PCR-based quantification

For qPCR analysis of barcode abundances, ligated fragments bound to beads were diluted 1:100 in water. 1.5 µl of beads at 0.1 µg/µl were added to 8.5 µl of qPCR master mix and amplified in StepOne™ Real-Time PCR system (Thermo Fisher Scientific Inc., USA).

7.3.2. Protocol for combinatorial barcoded RNA-Seq libraries

Purification and cDNA synthesis

This protocol describes the library preparation for one replicate of combinatorially barcoded mRNA of up to 420 combinations. Droplets containing lysed cells were incubated for 30 min at room temperature for the ligation of barcode fragments. 25 ml of 6xSSC was added to the droplets before they were broken by 1 ml of PFO. The solution was centrifuged for 2 min at 4000 g and 4 °C and the supernatant was transferred into a fresh 50 ml tube. The solution was again centrifuged at 4000 g and 4 °C for 5 min and the supernatant was transferred into a fresh 50 ml tube containing 250 µl of C1 dynabeads at 10 mg/ml. Barcodes were bound to the beads during a 20 min incubations step at RT under rotation. Beads were washed three times in 6xSSC using an in-house magnetic rack. Subsequently, beads were re-suspended in 400 µl RT mix and incubated under rotation for 30 min at room temperature followed for 90 min at 42 °C. The cDNA libraries were washed in TE-SDS and two times in water and then resuspened in 500 µl water. Barcoded cDNA libraries were cleaved off the beads using MseI (New England BioLabs Inc., USA) restriction digestion. Released fragments were sized selected by two rounds of purification using 0.6x the volume of Ampure XT beads and then amplified with the whole transcriptome amplification PCR program using KAPA Hifi ready mix (Kapa Biosystems, USA) with 0.8 µM SMART primer. PCR products were again purified on 0.6x the volume of Ampure XT beads followed by a fragment size analysis using high sensitivity DNA chips and a 2100 Bioanalyzer instrument (Agilent

Technologies Inc., USA). Samples were diluted to 0.2 ng/µl and used for tagmentation.

Tagmentation-based 3'-end NGS library preparation

Fragmentation of cDNA libraries and introduction of linker sequences was facilitated using Tn5-based tagmentation. We used a Tn5-based library preparation protocol that was developed in house (Hennig et al., 2018). First, linker oligonucleotides (Tn5ME-A and Tn5ME-B) were annealed with Tn5ME-Rev and mixed with a Tn5 stock (0.5 mg/ml) at 2.5 µM. For loading Tn5 with linkers, the mix was incubated for 45 min at 23 °C under constant shacking at 350 rpm. 8.25 µg/ml Tn5 with linker was mixed with 50 pg/µl cDNA and one volume of Tagmentation mix. The reaction mixture was incubated for 3 min at 55 °C and subsequently the Tn5 was inactivated by adding 0.2% SDS. Fragmented cDNA libraries were amplified using the P5-SMART adapter primer and i7 adapter index primers (Illumina Inc., USA) with the Tagmentation PCR Mix. Amplification products were purified by adding one volume of AMPure beads and size and concentration of the fragments were determined using a high sensitivity DNA chips in a 2100 Bioanalyzer. Subsequently, all samples were pooled at equimolar ratios and sequences on a NextSeq 500 (Illumina Inc., USA) machine together with PhiX spike-ins. Paired-end libraries of barcode combinations (Read 1, 26 bp) and cDNA (Read 2, 59 bp) were generated. In order to sequence the barcode combinations (Read 1), the custom sequencing primer was used.

7.4. Determining the growth inhibition rate of drugs

Drugs were tested for their effect on cell growth over three different concentration ranges:

1. 100µM, 10 µM, 1 µM
2. 10 µM, 1 µM, 0.1 µM
3. 1 µM, 0.1 µM and 0.01 µM

Concentration ranges for individual drugs were selected based on literature. K562 cells were suspended at 10^5 cell/ml in IMDM media (Thermo Fisher Scientific Inc., USA) supplemented with 10% FBS and 200 µl were pipetted in black clear bottom Corning 96-well plates (Sigma Aldrich, USA). Drugs were added and each drug

concentration was tested in triplicates. For t_0 measurements, 22 µl of PrestoBlue cell viability reagent (Thermo Fisher Scientific Inc., USA) was added immediately after the cells were plated and measured after 1 h of incubation. All other samples were incubated for 48 h before 22 µl of PrestoBlue was added to each well. Fluorescence signals were measured using a BioTek Synergy (BioTek Instruments Inc, VT) plate reader.

7.5. Data Analysis

7.5.1. Fluorescence measurements

Recorded intensities measured by the PMTs were analysed using R scripts. Fluorescence intensities of compound or sample plugs was plotted and analysed using the BraDiPlus package (Eduati et al., 2018). Fluorescence data recorded from droplets was visualized and analysed using ggpubr package tools. The fractions between standard deviation and mean was used to determine the coefficient of variance.

7.5.2. Determination of GR_{20} values

To compensate for cell division rates, fluorescence measurements of cell viability of treated and untreated K562 cell was analysed using the R package GRmetrics (Hafner et al., 2016). The growth rate inhibition (GR) is defined as the ratio of cell growth under drug treatment at a given concentration and cell growth under untreated conditions, which is normalized to the cell division rate. First, concentrations at which cell growth was inhibited by 50% (GR_{50}) and the Hill slope of the sigmoidal fit were determined. GR_{50} values and Hill slope (h) were used to calculate the GR_{20} values for each drug using the following equation:

$$GR_F = \left(\frac{F}{100 - F}\right)^{1/H} * GR_{50}$$

where F is the fraction of the maximal response which in this case was 20.

7.5.3. Demultiplexing of sequenced barcode libraries

In order to determine the abundance of barcode combinations generated in droplets, sequenced barcodes were first analysed using starcode (Zorita et al.,

2015) with default settings. Starcode clusters sequences according to their similarity (max. two mismatches) and reports read counts for each sequence. Starcode results were processed using a Python script that assigned barcode combination and counts to combinations of indices used for their identification, outputting a table with barcode identifiers sequences and read counts.

7.5.4. Barcode abundance

Read counts for individual replicates were first normalized by the library size. Barcode counts were visualized using the R package ggpubr. Fold changes between different collection times of barcodes were determined by dividing read counts by the mean count at a 5 sec collection for each set of seven combinations (A1-A7, B1-B7, etc.). Log2 of fractions was used to visualize barcode abundance and to statistically compare abundances between increments of droplet numbers using a pairwise Wilcox Rank Sum test. P-values were adjusted for multiple comparisons according to Benjamini & Hochberg (Benjamini and Hochberg, 1995).

7.5.5. Demultiplexing of sorted droplets by barcodes

Read counts from sorted and unsorted samples were normalized by their library size. Normalized read counts for each barcode were divided by the mean count counts of the corresponding barcode in the unsorted sample. The resulting fold-changes for each barcode was log2 transformed and p-values between sorted and unsorted barcodes were calculated using a Student's T-Test. Calculated p-values were corrected for multiple comparisons according to Benjamini & Hochberg. The overall enrichment of sorted barcodes and its statistical significance was determined by comparing log2 fold-changes of BC-3A to BC-3G to log2-fold changes of all other barcodes using a Student's T-Test.

7.5.6. Demultiplexing and alignments of RNA-Seq data

Paired-end libraries of combinatorial barcoded mRNA was first demultiplexed using the in-line i7 Illumina indices to assign all reads to replicates. In the next step, paired-end libraries were demultiplex based on sequenced barcode combinations. Demultiplexing was performed using the Je demultiplexing software for NGS data (Girardot et al., 2016). Sequenced barcodes (Read 1) were used to identify the treatment conditions by comparing Read 1 to a reference of known barcode (each

assigned to a drug combination). Up to 6 mismatches between Read 1 and the known barcode were allowed to unambiguously assign the sequenced mRNA (Read 2) to each drug combination. Once mRNA sequences were assigned to barcode combinations (i.e. drug treatments), the data was aligned to a reference genome using the STAR RNA-Seq alignment tool (Dobin et al., 2013). Demultiplexing and aligning of sequencing results was done by Jonathan Landry from the EMBL Genomic Core Facility.

7.5.7. Filtering, normalization and transformation of RNA-Seq data

Demultiplexed and aligned data was analysed using Limma and EdgeR packages (Ritchie et al., 2015; Robinson et al., 2010). Read count tables obtained form STAR were loaded into R to create a read-count matrix comprising all samples. First, samples were filtered based on library size by removing samples exceeding one standard deviation from the mean library size (mean +/- sd). In a next step, we filtered out genes that were not expression in 10% or 50% of all samples followed by trimmed mean of M-values (TMM) normalization: Scale factors for each sample were calculated and used to determine the effective library sizes for individual samples. Subsequently, the data was transformed using voom and corrected for batch effects by calling the corresponding functions in the Limma package (Ritchie et al., 2015). Dimensional reduction of transformed data sets was performed using t-distributed stochastic neighbourhood embedding (tSNE, Rtsne) with a perplexity of 30. Gene expression signatures were determined by calculating z-scores of gene expression for each sample and used to calculate spearman correlations between samples.

7.6. Drugs

Drug	Targets	ChEMBL ID	Clinical Phase	Provider	Drug Class
10Z-Hymenialdisine	MEK1	CHEMBL361708	0	TOCRIS	Kinase-Inhibitor
5-Fluorouracil	TYMS	CHEMBL185	4		Anti-metabolite
5-Iodotubercidine	ADK/INSR/ PKA/CK1	CHEMBL99203	0	Selleck	Anti-metabolite
AT9283	AURKA/AURKB/JAK2/JAK3	CHEMBL495727	2	Selleck	Kinase-Inhibitor
Baricitinib	JAK1/JAK2	CHEMBL2105759	4	Selleck	Kinase-Inhibitor
Blebbistatin	MYH2	CHEMBL1328324	0	Selleck	Cytoskelleton
Clofarabine	RRM1 / DNA Polymerase	CHEMBL1750	4	Selleck	Anti-metabolite
Cytarabine	DNA / RNA Polymerase	CHEMBL803	4	Selleck	Anti-metabolite
Dacarbazine	DNA Animetabolite	CHEMBL476	4	Selleck	Alkylating Agent
Decitabine	DNMT1	CHEMBL1201129	4	Selleck	Anti-metabolite
Dexrazoxane	TOP2	CHEMBL1738	4	Selleck	Anthracycline
Dovitinib	FLT3/c-Kit/FGFR1/FGFR3	CHEMBL522892	3	Selleck	Kinase-Inhibitor
Doxorubicin	TOP2	CHEMBL53463	4	Selleck	Anthracycline
EPIRUBICIN	TOP2	CHEMBL1200981	4	Selleck	Anthracycline
Fludarabine Phosphate	DNA Animetabolite	CHEMBL1096882	4	Selleck	Anti-metabolite
Gemcitabine	DNA Animetabolite	CHEMBL888	4	Selleck	Anti-metabolite
Gimeracil	DPYD	CHEMBL1730601	3	Selleck	
H-7 dihydrochloride	PRKC/PKG/PKA		0	TOCRIS	Kinase-Inhibitor
Hematoxylin	EGFR/ERBB2/c-MET/c-KIT/SRC	CHEMBL477197	0	Selleck	
Imatinib	BCR-ABL	CHEMBL941	4	Selleck	Kinase-Inhibitor

Methodextrat	DHFR	CHEMBL34259	4	Selleck	Kinase-Inhibitor
Mitomycin C	DNA Synthesis	CHEMBL105	4	Selleck	DNA-crosslinker
Nelarabine	DNA Animetabolite	CHEMBL1201112	4	Selleck	Anti-Metabolite
NMS-1286937	PLK1	CHEMBL1094408	1	Selleck	Kinase-Inhibitor
Olomoucine	CDK2/MAPK3	CHEMBL280074	0	TOCRIS	Kinase-Inhibitor
Oxaliplatin	DNA	CHEMBL414804	4	Selleck	Alkylating Agent
PF-562271	FAK	CHEMBL1084546	1		
Pomalidomide	TNF-alpha	CHEMBL43452	4	Selleck	
Sangivamycin	PKC	CHEMBL101892	0	Santa Cruz	
SB-747651	MSK1/MSK2	CHEMBL188434	0	TOCRIS	Kinase-Inhibitor
SF-1126	PI3K/mTOR	CHEMBL2326966	2	Santa Cruz	Kinase-Inhibitor
Sonolisib	PI3K	CHEMBL411907	2	Abcam	Kinase-Inhibitor
Streptozotocin	DNA	CHEMBL1603	0	Selleck	
Sunitinib Malate	VEGFR2/PDGFRb	CHEMBL1567	4	Selleck	Kinase-Inhibitor
Tabloid / Thioguanine	DNMT1	CHEMBL727	4	Selleck	Anti-metabolite
Thiostrepton	FOXM1			Santa Cruz	
Trametinib	MEK1/2	CHEMBL2103875	4	Selleck	Kinase-Inhibitor
Triciribine	AKT1/AKT2/AKT3	CHEMBL462018	2	Selleck	Kinase-Inhibitor
Wortmannin	PI3K	CHEMBL428496	0	Selleck	Kinase-Inhibitor
YM155	BIRC5	CHEMBL2110734	2	Selleck	

7.7. Buffers and solutions

3x Ligation Mix	For 200 µl: 30 µl T4 DNA Ligase (New England Biolabs, USA) 60 µl 10x T4 DNA Ligation Buffer (New England Biolabs, USA) 110 µl H_2O (Thermo Fisher Inc., USA)
Ligation and Lysis Mix	For 600 µl: 90 µl T4 DNA Ligase (New England Biolabs, USA) 90 µl NxGen RNase Inhibitor (Lucigen, USA) 54 µl 10% Igepal (Sigma Aldrich, USA) 180 µl 10x T4 DNA Ligation Buffer (New England Biolabs, USA) 186 µl H_2O (Thermo Fisher Inc., USA)
TE-SDS	10 mM Tris pH 8 1 mM EDTA 0.5% SDS
Phusion PCR Mix	1 x Reaction (15 µl): 3 µl 5x Phusion GC Buffer (New England Biolabs, USA) 0.4 µl 10 mM dNTPs (Clontech, USA) 0.45 µl DMSO (New England Biolabs, USA) 0.15 µl Phusion DNA Polymerase (New England Biolabs, USA) 5 µl H_2O (Thermo Fisher Inc., USA) 1 µl NEXTFlex Amplicon-Seq primer mix (Bioo Scientific Corp.) 5 µl Barcode library
qPCR master mix	1 x Reaction (10 µl) 5 µl SYBR™ Green PCR Master Mix 0.5 µl 10 µl Fwd. Primer 0.5 µl 10 µl Rev Primer 2.5 µl H_2O (Thermo Fisher Inc., USA) 1.5 µl sample (barcode combination)
RT mix	For 400 µl: 150 µl H_2O 80 µl 10x RT Buffer (Thermo Fisher Scientific Inc.) 80 µl 20% Ficoll-PM 400 (Sigma Aldrich) 40 µl 10 mM dNTPs (Clontech) 10 µl NXGen RNase Inhibitor (Lucigen, USA) 20 µl 50 µM TSO 20 µl Maxima -H RT (Thermo Fisher Scientific Inc.)

Tagmentation Mix	20 mM Tris-HCL pH 7.5 20 mM $MgCl_2$ 50 % v/v 100% DMF (Sigma Aldrich, USA)
Tagmentation PCR Mix	1 x Reaction (30 µl) 20.25 µl KAPA Hifi Ready mix (Kapa Biosystems, USA) 2.25 µl DMSO (Sigma Aldrich, USA) 3.75 µl 10 µM P5-SMART adapter primer 3. 75 µl 10 µM i7 adapter index primer (Illumina, USA) + 18.75 µl Tagmentation products

7.8. PCR programs

Phusion PCR for barcode amplification		
Step	Temperature	Time
Initial denaturation	98 °C	30 sec
10 cycles	98 °C	10 sec
	70 °C	20 sec
	72 °C	20 sec
Hold	4 °C	

Whole transcriptome amplification		
Step	Temperature	Time
Initial denaturation	95 °C	3 min
4 cycles	98 °C	20 sec
	65 °C	45 sec
	72 °C	3 min
9 cycles	98 °C	20 sec
	67 °C	20 sec
	72 °C	3 min
Final extension	72	5 min
Hold	4 °C	

Tagmentation PCR		
Step	Temperature	Time
Initial denaturation	95 °C	30 sec
12 cycles	98 °C	20 sec
	58 °C	15 sec
	72 °C	30 sec
Final extension	72	3 min
Hold	10 °C	

7.9. DNA sequences

Sequence Name	Sequence (5’ -> 3’)
TSO	AAGCAGTGGTATCAACGCAGAGTGAATrGrGrG
SMART-Primer	AAGCAGTGGTATCAACGCAGAGT
Tn5ME-A	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
Tn5ME-B	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
Tn5MErev	[phos]CTGTCTCTTATACACATCT
i7 index adapter primer	CAAGCAGAAGACGGCATACGAGATnnnnnnnnGTCTCGTGGGCTCGG
P5- P5-SMART adapter primer	AATGATACGGCGACCACCGAGATCTACACGCCT GTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC
Custom Sequencing Primer	GCCTGTCCGCGGAAGCAGTGGTATCAACGCAG AGTAC

7.9.1. Barcode sequences for RNA-Seq

Sequences BC-bio

BC-1	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACGCTCACCTGCgcggc
	[Pho]GCAGGTGAGCGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-2	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACTA CCGTCTCGgcggc
	[Pho]CGAGACGGTAGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-3	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTAC AAGCTGGCTCgcggc
	[Pho]GAGCCAGCTTGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-4	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACAAGTATCCTCgcggc
	[Pho]GAGGATACTTGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-5	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACATCAATCCTCgcggc
	[Pho]GAGGATTGATGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-6	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACACTCACCCTCgcggc
	[Pho]GAGGGTGAGTGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-7	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACTGACGTACTCgcggc
	[Pho]GAGTACGTCAGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-8	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACGTCAAGCATCgcggc
	[Pho]GATGCTTGACGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-9	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACTGCCGTCAAGgcggc
	[Pho]CTTGACGGCAGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-10	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACAATTCCCAGGgcggc
	[Pho]CCTGGGAATTGTACTCTGCGTTGATACCACTGCTTAAAAAAA

BC-11	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACCGTCTAACGAgcggc
	[Pho]TCGTTAGACGGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-12	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACCCAATATGTGgcggc
	[Pho]CACATATTGGGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-13	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACGCGTTATGCAgcggc
	[Pho]TGCATAACGCGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-14	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACACTACGTAGTgcggc
	[Pho]ACTACGTAGTGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-15	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACGGCCAATCATgcggc
	[Pho]ATGATTGGCCGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-16	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACTGTCAACAGGgcggc
	[PHO]CCTGTTGACAGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-17	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACATACCGTTGCgcggc
	[PHO]GCAACGGTATGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-18	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACCCAGACATTGgcggc
	[PHO]CAATGTCTGGGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-19	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACTGTACAGGCTgcggc
	[PHO]AGCCTGTACAGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-20	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACATGTGCATGCgcggc
	[PHO]GCATGCACATGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-21	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACGACAGATTCGgcggc
	[PHO]CGAATCTGTCGTACTCTGCGTTGATACCACTGCTTAAAAAAA
BC-22	[BIO]TTTTTTTAAGCAGTGGTATCAACGCAGAGTACTCACCGAGATgcggc
	[PHO]ATCTCGGTGAGTACTCTGCGTTGATACCACTGCTTAAAAAAA

Sequences BC-dT

BC-A	[Pho]GCTGTGACTGTTTTTTTTTTTTTTTTTTTVN
	CAGTCACAGCgccgc
BC-B	[Pho]GGACGCA TAGTTTTTTTTTTTTTTTTTTTVN
	CTATGCGTCCgccgc
BC-C	[Pho]GGAGGATTCCTTTTTTTTTTTTTTTTTTTVN
	GGAATCCTCCgccgc
BC-D	[Pho]ACATATATCCTTTTTTTTTTTTTTTTTTTVN
	GGATATATGTgccgc
BC-E	[Pho]GGATGGAACATTTTTTTTTTTTTTTTTTTVN
	TGTTCCATCCgccgc
BC-F	[Pho]ACTCTTCGTGTTTTTTTTTTTTTTTTTTTVN
	CACGAAGAGTgccgc
BC-G	[Pho]CAGTTCCACCTTTTTTTTTTTTTTTTTTTVN

	GGTGGAACTGgccgc
BC-H	[Pho]CCAGGGAATTTTTTTTTTTTTTTTTTTTVN
	CATTCCCTGGgccgc
BC-I	[Pho]CGGTACTTACTTTTTTTTTTTTTTTTTTVN
	GTAAGTACCGgccgc
BC-J	[Pho]TCACTGATAGTTTTTTTTTTTTTTTTTTVN
	CTATCAGTGAgccgc
BC-K	[Pho]TATCGCGGACTTTTTTTTTTTTTTTTTTVN
	GTCCGCGATAgccgc
BC-L	[Pho]CAGAGTGCCTTTTTTTTTTTTTTTTTTVN
	AGGCACTCTGgccgc
BC-M	[Pho]AGATTCGGCTTTTTTTTTTTTTTTTTTVN
	AGCCGAATCTgccgc
BC-N	[Pho]GTCACGGTCATTTTTTTTTTTTTTTTTTVN
	TGACCGTGACgccgc
BC-O	[Pho]ACAGTTGCTGTTTTTTTTTTTTTTTTTTVN
	CAGCAACTGTgccgc
BC-P	[PHO]ATGTAGCACCTTTTTTTTTTTTTTTTTTVN
	GGTGCTACATgccgc
BC-Q	[PHO]CCGTACAAGTTTTTTTTTTTTTTTTTTVN
	ACTTGTACGGgccgc
BC-R	[PHO]CGTAGTCCAGTTTTTTTTTTTTTTTTTTVN
	CTGGACTACGgccgc
BC-S	[PHO]AACGCTGTAGTTTTTTTTTTTTTTTTTTVN
	CTACAGCGTTgccgc
BC-T	[PHO]GAGGTTCCCATTTTTTTTTTTTTTTTTTVN
	TGGGAACCTCgccgc
BC-U	[PHO]GTCCGGAACTTTTTTTTTTTTTTTTTTVN
	AGTTCCGGACgccgc
BC-V	[PHO]TACATCGCAGTTTTTTTTTTTTTTTTTTVN
	CTGCGATGTAgccgc

7.9.2. Sequences for droplet barcoding

BC-1	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTGCTCACCTGCgcggc
	[Pho]GCAGGTGAGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAG
BC-2	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTTACCGTCTCGgcggc
	[Pho]CGAGACGGTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAG

BC-3	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTAAGCTGGCTCgcggc
	[Pho]GAGCCAGCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAG
BC-4	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTAAGTATCCTCgcggc
	[Pho]GAGGATACTT AGATCGGAAGAGCGTCGTGTAGGGAAAGAG
BC-5	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTATCAATCCTCgcggc
	[Pho]GAGGATTGAT AGATCGGAAGAGCGTCGTGTAGGGAAAGAG
BC-6	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTACTCACCCTCgcggc
	[Pho]GAGGGTGAGT AGATCGGAAGAGCGTCGTGTAGGGAAAGAG
BC-7	[Bio]CTCTTTCCCTACACGACGCTCTTCCGATCTTGACGTACTCgcggc
	[Pho]GAGTACGTCA AGATCGGAAGAGCGTCGTGTAGGGAAAGAG

BC-A	CTGGAGTTCAGACGTGTGCTCTTCCGATCT CAGTCACAGCgccgc
	[Pho]GCTGTGACTG AGATCGGAAGAGCACACGTCTGAACTCCAG
BC-B	[Pho]GGACGCATAG AGATCGGAAGAGCACACGTCTGAACTCCAG
	CTGGAGTTCAGACGTGTGCTCTTCCGATCT CTATGCGTCCgccgc
BC-C	[Pho]GGAGGATTCC AGATCGGAAGAGCACACGTCTGAACTCCAG
	CTGGAGTTCAGACGTGTGCTCTTCCGATCT GGAATCCTCCgccgc
BC-D	[Pho]ACATATATCC AGATCGGAAGAGCACACGTCTGAACTCCAG
	CTGGAGTTCAGACGTGTGCTCTTCCGATCT GGATATATGTgccgc
BC-E	[Pho]GGATGGAACA AGATCGGAAGAGCACACGTCTGAACTCCAG
	CTGGAGTTCAGACGTGTGCTCTTCCGATCT TGTTCCATCCgccgc
BC-F	[Pho]ACTCTTCGTG AGATCGGAAGAGCACACGTCTGAACTCCAG
	CTGGAGTTCAGACGTGTGCTCTTCCGATCT CACGAAGAGTgccgc
BC-G	[Pho]CAGTTCCACC AGATCGGAAGAGCACACGTCTGAACTCCAG
	CTGGAGTTCAGACGTGTGCTCTTCCGATCT GGTGGAACTGgccgc

8. References

Abate, A.R., Hung, T., Mary, P., Agresti, J.J., and Weitz, D.A. (2010). High-throughput injection with microfluidics using picoinjectors. Proc Natl Acad Sci U S A *107*, 19163-19166.

Abate, A.R., Kutsovsky, M., Seiffert, S., Windbergs, M., Pinto, L.F., Rotem, A., Utada, A.S., and Weitz, D.A. (2011). Synthesis of monodisperse microparticles from non-Newtonian polymer solutions with microfluidic devices. Adv Mater *23*, 1757-1760.

Agresti, J.J., Antipov, E., Abate, A.R., Ahn, K., Rowat, A.C., Baret, J.C., Marquez, M., Klibanov, A.M., Griffiths, A.D., and Weitz, D.A. (2010). Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. Proc Natl Acad Sci U S A *107*, 4004-4009.

Al-Lazikani, B., Banerji, U., and Workman, P. (2012). Combinatorial drug therapy for cancer in the post-genomic era. Nat Biotechnol *30*, 679-691.

Alberts, B. (2015). Molecular biology of the cell, Vol Sixth Edition (New York, NY: Garland Science).

Anna, S.L., Bontoux, N., and Stone, H.A. (2003). Formation of dispersions using "flow focusing" in microchannels. Appl Phys Lett *82*, 364-366.

Bansal, M., Yang, J., Karan, C., Menden, M.P., Costello, J.C., Tang, H., Xiao, G., Li, Y., Allen, J., Zhong, R.*, et al.* (2014). A community computational challenge to predict the activity of pairs of compounds. Nat Biotechnol *32*, 1213-1222.

Baret, J.C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., Rick, C., Samuels, M.L., Hutchison, J.B., Agresti, J.J.*, et al.* (2009). Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. Lab Chip *9*, 1850-1858.

Barnes, D.J., Palaiologou, D., Panousopoulou, E., Schultheis, B., Yong, A.S., Wong, A., Pattacini, L., Goldman, J.M., and Melo, J.V. (2005). Bcr-Abl expression levels determine the rate of development of resistance to imatinib mesylate in chronic myeloid leukemia. Cancer Res *65*, 8912-8919.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) *57*, 289-300.

Bild, A.H., Yao, G., Chang, J.T., Wang, Q., Potti, A., Chasse, D., Joshi, M.B., Harpole, D., Lancaster, J.M., Berchuck, A., *et al.* (2006). Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature *439*, 353-357.

Borisy, A.A., Elliott, P.J., Hurst, N.W., Lee, M.S., Lehar, J., Price, E.R., Serbedzija, G., Zimmermann, G.R., Foley, M.A., Stockwell, B.R., *et al.* (2003). Systematic discovery of multicomponent therapeutics. Proc Natl Acad Sci U S A *100*, 7977-7982.

Bremond, N., Thiam, A.R., and Bibette, J. (2008). Decompressing emulsion droplets favors coalescence. Phys Rev Lett *100*, 024501.

Brochado, A.R., Telzerow, A., Bobonis, J., Banzhaf, M., Mateus, A., Selkrig, J., Huth, E., Bassler, S., Zamarreno Beas, J., Zietek, M., *et al.* (2018). Species-specific activity of antibacterial drug combinations. Nature *559*, 259-263.

Brouzes, E., Medkova, M., Savenelli, N., Marran, D., Twardowski, M., Hutchison, J.B., Rothberg, J.M., Link, D.R., Perrimon, N., and Samuels, M.L. (2009). Droplet microfluidic technology for single-cell high-throughput screening. Proc Natl Acad Sci U S A *106*, 14195-14200.

Burrell, R.A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. Nature *501*, 338-345.

Bush, E.C., Ray, F., Alvarez, M.J., Realubit, R., Li, H., Karan, C., Califano, A., and Sims, P.A. (2017). PLATE-Seq for genome-wide regulatory network analysis of high-throughput screens. Nat Commun *8*, 105.

Choudhury, D., van Noort, D., Iliescu, C., Zheng, B., Poon, K.L., Korzh, S., Korzh, V., and Yu, H. (2012). Fish and Chips: a microfluidic perfusion platform for monitoring zebrafish development. Lab Chip *12*, 892-900.

Clausell-Tormos, J., Griffiths, A.D., and Merten, C.A. (2010). An automated two-phase microfluidic system for kinetic analyses and the screening of compound libraries. Lab Chip *10*, 1302-1307.

Clausell-Tormos, J., Lieber, D., Baret, J.C., El-Harrak, A., Miller, O.J., Frenz, L., Blouwolff, J., Humphry, K.J., Koster, S., Duan, H., *et al.* (2008). Droplet-based microfluidic platforms for the encapsulation and screening of Mammalian cells and multicellular organisms. Chem Biol *15*, 427-437.

Crick, F.H. (1958). On protein synthesis. Symp Soc Exp Biol *12*, 138-163.

Croce, C.M. (2008). Oncogenes and cancer. N Engl J Med *358*, 502-511.

Crystal, A.S., Shaw, A.T., Sequist, L.V., Friboulet, L., Niederst, M.J., Lockerman, E.L., Frias, R.L., Gainor, J.F., Amzallag, A., Greninger, P., *et al.* (2014). Patient-derived models of acquired resistance can identify effective drug combinations for cancer. Science *346*, 1480-1486.

Damia, G., and D'Incalci, M. (1998). Mechanisms of resistance to alkylating agents. Cytotechnology *27*, 165-173.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W.*, et al.* (2002). Mutations of the BRAF gene in human cancer. Nature *417*, 949-954.

De Palma, M., and Hanahan, D. (2012). The biology of personalized cancer medicine: facing individual complexities underlying hallmark capabilities. Mol Oncol *6*, 111-127.

Dietlein, F., Kalb, B., Jokic, M., Noll, E.M., Strong, A., Tharun, L., Ozretic, L., Kunstlinger, H., Kambartel, K., Randerath, W.J.*, et al.* (2015). A Synergistic Interaction between Chk1- and MK2 Inhibitors in KRAS-Mutant Cancer. Cell *162*, 146-159.

Dietrich, S., Oles, M., Lu, J., Sellner, L., Anders, S., Velten, B., Wu, B., Hullein, J., da Silva Liberio, M., Walther, T.*, et al.* (2018). Drug-perturbation-based stratification of blood cancer. J Clin Invest *128*, 427-445.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Dovichi, N.J., and Zhang, J. (2000). How Capillary Electrophoresis Sequenced the Human Genome This Essay is based on a lecture given at the Analytica 2000 conference in Munich (Germany) on the occasion of the Heinrich-Emanuel-Merck Prize presentation. Angew Chem Int Ed Engl *39*, 4463-4468.

Druker, B.J. (2008). Translation of the Philadelphia chromosome into therapy for CML. Blood *112*, 4808-4817.

Duffy, D.C., McDonald, J.C., Schueller, O.J., and Whitesides, G.M. (1998). Rapid Prototyping of Microfluidic Systems in Poly(dimethylsiloxane). Anal Chem *70*, 4974-4984.

Eduati, F., Utharala, R., Madhavan, D., Neumann, U.P., Longerich, T., Cramer, T., Saez-Rodriguez, J., and Merten, C.A. (2018). A microfluidics platform for combinatorial drug screening on cancer biopsies. Nat Commun *9*, 2434.

Eicher, D., and Merten, C.A. (2011). Microfluidic devices for diagnostic applications. Expert Rev Mol Diagn *11*, 505-519.

El Debs, B., Utharala, R., Balyasnikova, I.V., Griffiths, A.D., and Merten, C.A. (2012). Functional single-cell hybridoma screening using droplet-based microfluidics. Proc Natl Acad Sci U S A *109*, 11570-11575.

Engelman, J.A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J.O., Lindeman, N., Gale, C.M., Zhao, X., Christensen, J.*, et al.* (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science *316*, 1039-1043.

Eyer, K., Doineau, R.C.L., Castrillon, C.E., Briseno-Roa, L., Menrath, V., Mottet, G., England, P., Godina, A., Brient-Litzler, E., Nizak, C.*, et al.* (2017). Single-cell deep

phenotyping of IgG-secreting cells for high-resolution immune monitoring. Nat Biotechnol *35*, 977-982.

Fan, H.C., Wang, J., Potanina, A., and Quake, S.R. (2011). Whole-genome molecular haplotyping of single cells. Nat Biotechnol *29*, 51-57.

Feuk, L., Carson, A.R., and Scherer, S.W. (2006). Structural variation in the human genome. Nat Rev Genet *7*, 85-97.

Frisch, S.M., and Screaton, R.A. (2001). Anoikis mechanisms. Curr Opin Cell Biol *13*, 555-562.

Fu, Y., Li, C., Lu, S., Zhou, W., Tang, F., Xie, X.S., and Huang, Y. (2015). Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. Proc Natl Acad Sci U S A *112*, 11923-11928.

Gao, H., Korn, J.M., Ferretti, S., Monahan, J.E., Wang, Y., Singh, M., Zhang, C., Schnell, C., Yang, G., Zhang, Y.*, et al.* (2015). High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. Nat Med *21*, 1318-1325.

Genomes Project, C., Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.

Ghose, A.K., and Crippen, G.M. (1987). Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. Journal of Chemical Information and Computer Sciences *27*, 21-35.

Giaever, I., and Keese, C.R. (1983). Behavior of cells at fluid interfaces. Proc Natl Acad Sci U S A *80*, 219-222.

Gielen, F., Buryska, T., Van Vliet, L., Butz, M., Damborsky, J., Prokop, Z., and Hollfelder, F. (2015). Interfacing microwells with nanoliter compartments: a sampler generating high-resolution concentration gradients for quantitative biochemical analyses in droplets. Anal Chem *87*, 624-632.

Girardot, C., Scholtalbers, J., Sauer, S., Su, S.Y., and Furlong, E.E. (2016). Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. BMC Bioinformatics *17*, 419.

Gottesman, M.M., Fojo, T., and Bates, S.E. (2002). Multidrug resistance in cancer: role of ATP-dependent transporters. Nat Rev Cancer *2*, 48-58.

Gruner, P., Riechers, B., Semin, B., Lim, J., Johnston, A., Short, K., and Baret, J.C. (2016). Controlling molecular transport in minimal emulsions. Nat Commun *7*, 10392.

Gulick, R.M., Mellors, J.W., Havlir, D., Eron, J.J., Gonzalez, C., McMahon, D., Richman, D.D., Valentine, F.T., Jonas, L., Meibohm, A.*, et al.* (1997). Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. N Engl J Med *337*, 734-739.

Hackam, D.G., and Redelmeier, D.A. (2006). Translation of research evidence from animals to humans. JAMA *296*, 1731-1732.

Haessler, U., Pisano, M., Wu, M., and Swartz, M.A. (2011). Dendritic cell chemotaxis in 3D under defined chemokine gradients reveals differential response to ligands CCL21 and CCL19. Proc Natl Acad Sci U S A *108*, 5614-5619.

Hafner, M., Niepel, M., Chung, M., and Sorger, P.K. (2016). Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. Nat Methods *13*, 521-527.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57-70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646-674.

Hassane, D.C., Sen, S., Minhajuddin, M., Rossi, R.M., Corbett, C.A., Balys, M., Wei, L., Crooks, P.A., Guzman, M.L., and Jordan, C.T. (2010). Chemical genomic screening reveals synergism between parthenolide and inhibitors of the PI-3 kinase and mTOR pathways. Blood *116*, 5983-5990.

Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.E., Krishnamurthy Radhakrishna, V.*, et al.* (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat Med *22*, 262-269.

He, L., Tang, J., Andersson, E.I., Timonen, S., Koschmieder, S., Wennerberg, K., Mustjoki, S., and Aittokallio, T. (2018). Patient-Customized Drug Combination Prediction and Testing for T-cell Prolymphocytic Leukemia Patients. Cancer Res *78*, 2407-2418.

Hennig, B.P., Velten, L., Racke, I., Tu, C.S., Thoms, M., Rybin, V., Besir, H., Remans, K., and Steinmetz, L.M. (2018). Large-Scale Low-Cost NGS Library Preparation Using a Robust Tn5 Purification and Tagmentation Protocol. G3 (Bethesda) *8*, 79-89.

Holtze, C., Rowat, A.C., Agresti, J.J., Hutchison, J.B., Angile, F.E., Schmitz, C.H., Koster, S., Duan, H., Humphry, K.J., Scanga, R.A.*, et al.* (2008). Biocompatible surfactants for water-in-fluorocarbon emulsions. Lab Chip *8*, 1632-1639.

Huberts, D.H., Sik Lee, S., Gonzales, J., Janssens, G.E., Vizcarra, I.A., and Heinemann, M. (2013). Construction and use of a microfluidic dissection platform for long-term imaging of cellular processes in budding yeast. Nat Protoc *8*, 1019-1027.

Hulett, H.R., Bonner, W.A., Barrett, J., and Herzenberg, L.A. (1969). Cell sorting: automated separation of mammalian cells as a function of intracellular fluorescence. Science *166*, 747-749.

Hwang, B., Lee, J.H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med *50*, 96.

Iorio, F., Bosotti, R., Scacheri, E., Belcastro, V., Mithbaokar, P., Ferriero, R., Murino, L., Tagliaferri, R., Brunetti-Pierri, N., Isacchi, A.*, et al.* (2010). Discovery of drug mode

of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci U S A *107*, 14621-14626.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Goncalves, E., Barthorpe, S., Lightfoot, H.*, et al.* (2016). A Landscape of Pharmacogenomic Interactions in Cancer. Cell *166*, 740-754.

Iorio, F., Rittman, T., Ge, H., Menden, M., and Saez-Rodriguez, J. (2013). Transcriptional data: a new gateway to drug repositioning? Drug Discov Today *18*, 350-357.

Isakova, A., Groux, R., Imbeault, M., Rainer, P., Alpern, D., Dainese, R., Ambrosini, G., Trono, D., Bucher, P., and Deplancke, B. (2017). SMiLE-seq identifies binding motifs of single and dimeric transcription factors. Nat Methods *14*, 316-322.

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lonnerberg, P., and Linnarsson, S. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods *11*, 163-166.

Jia, J., Zhu, F., Ma, X., Cao, Z., Cao, Z.W., Li, Y., Li, Y.X., and Chen, Y.Z. (2009). Mechanisms of drug combinations: interaction and network perspectives. Nat Rev Drug Discov *8*, 111-128.

Keiser, M.J., Setola, V., Irwin, J.J., Laggner, C., Abbas, A.I., Hufeisen, S.J., Jensen, N.H., Kuijer, M.B., Matos, R.C., Tran, T.B.*, et al.* (2009). Predicting new molecular targets for known drugs. Nature *462*, 175-181.

Kim, C., Gao, R., Sei, E., Brandt, R., Hartman, J., Hatschek, T., Crosetto, N., Foukakis, T., and Navin, N.E. (2018). Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. Cell *173*, 879-893 e813.

Kim, S., Kim, H.J., and Jeon, N.L. (2010). Biological applications of microfluidic gradient devices. Integr Biol (Camb) *2*, 584-603.

Kirschner, K., and Melton, D.W. (2010). Multiple roles of the ERCC1-XPF endonuclease in DNA repair and resistance to anticancer drugs. Anticancer Res *30*, 3223-3232.

Klaeger, S., Heinzlmeir, S., Wilhelm, M., Polzer, H., Vick, B., Koenig, P.A., Reinecke, M., Ruprecht, B., Petzoldt, S., Meng, C.*, et al.* (2017). The target landscape of clinical kinase drugs. Science *358*.

Klein, A.M., and Macosko, E. (2017). InDrops and Drop-seq technologies for single-cell sequencing. Lab Chip *17*, 2540-2541.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187-1201.

Kobayashi, S., Boggon, T.J., Dayaram, T., Janne, P.A., Kocher, O., Meyerson, M., Johnson, B.E., Eck, M.J., Tenen, D.G., and Halmos, B. (2005). EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med *352*, 786-792.

Kola, I., and Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov *3*, 711-715.

Kulesa, A., Kehe, J., Hurtado, J.E., Tawde, P., and Blainey, P.C. (2018). Combinatorial drug discovery in nanoliter droplets. Proc Natl Acad Sci U S A *115*, 6685-6690.

Lan, F., Demaree, B., Ahmed, N., and Abate, A.R. (2017). Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding. Nat Biotechnol *35*, 640-646.

Lan, F., Haliburton, J.R., Yuan, A., and Abate, A.R. (2016). Droplet barcoding for massively parallel single-molecule deep sequencing. Nat Commun *7*, 11784.

Lin, M., Whitmire, S., Chen, J., Farrel, A., Shi, X., and Guo, J.T. (2017). Effects of short indels on protein structure and function in human genomes. Sci Rep *7*, 9313.

Lonostro, P. (1995). Phase-Separation Properties of Fluorocarbons, Hydrocarbons and Their Copolymers. Adv Colloid Interfac *56*, 245-287.

Lowe, K.C., Davey, M.R., and Power, J.B. (1998). Perfluorochemicals: their applications and benefits to cell culture. Trends Biotechnol *16*, 272-277.

Lucchetta, E.M., Lee, J.H., Fu, L.A., Patel, N.H., and Ismagilov, R.F. (2005). Dynamics of Drosophila embryonic patterning network perturbed in space and time using microfluidics. Nature *434*, 1134-1138.

Luni, C., Giulitti, S., Serena, E., Ferrari, L., Zambon, A., Gagliano, O., Giobbe, G.G., Michielin, F., Knobel, S., Bosio, A.*, et al.* (2016). High-efficiency cellular reprogramming with microfluidics. Nat Methods *13*, 446-452.

Ma, C., Fan, R., Ahmad, H., Shi, Q., Comin-Anduix, B., Chodon, T., Koya, R.C., Liu, C.C., Kwong, G.A., Radu, C.G.*, et al.* (2011). A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. Nat Med *17*, 738-743.

Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V., Hertzberg, R.P., Janzen, W.P., Paslay, J.W.*, et al.* (2011). Impact of high-throughput screening in biomedical research. Nat Rev Drug Discov *10*, 188-195.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M.*, et al.* (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202-1214.

Manchado, E., Weissmueller, S., Morris, J.P.t., Chen, C.C., Wullenkord, R., Lujambio, A., de Stanchina, E., Poirier, J.T., Gainor, J.F., Corcoran, R.B.*, et al.* (2016). A combinatorial strategy for treating KRAS-mutant lung cancer. Nature *534*, 647-651.

Mazutis, L., Baret, J.C., and Griffiths, A.D. (2009). A fast and efficient microfluidic system for highly selective one-to-one droplet fusion. Lab Chip *9*, 2665-2672.

Menden, M.P., Wang, D., Guan, Y., Mason, M., Szalai, B., Bulusu, K.C., Yu, T., Kang, J., Jeon, M., Wolfinger, R.*, et al.* (2018). A cancer pharmacogenomic screen powering crowd-sourced advancement of drug combination prediction. bioRxiv.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nat Rev Genet *11*, 31-46.

Miller, O.J., El Harrak, A., Mangeat, T., Baret, J.C., Frenz, L., El Debs, B., Mayot, E., Samuels, M.L., Rooney, E.K., Dieu, P.*, et al.* (2012). High-resolution dose-response screening using droplet-based microfluidics. Proc Natl Acad Sci U S A *109*, 378-383.

Montero, J., Sarosiek, K.A., DeAngelo, J.D., Maertens, O., Ryan, J., Ercan, D., Piao, H., Horowitz, N.S., Berkowitz, R.S., Matulonis, U.*, et al.* (2015). Drug-induced death signaling strategy rapidly predicts cancer response to chemotherapy. Cell *160*, 977-989.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D.*, et al.* (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90-94.

Negrini, S., Gorgoulis, V.G., and Halazonetis, T.D. (2010). Genomic instability--an evolving hallmark of cancer. Nat Rev Mol Cell Biol *11*, 220-228.

Nicolini, A., Ferrari, P., and Duffy, M.J. (2018). Prognostic and predictive biomarkers in breast cancer: Past, present and future. Semin Cancer Biol *52*, 56-73.

Pan, M., Lyu, F., and Tang, S.K. (2015). Fluorinated Pickering Emulsions with Nonadsorbing Interfaces for Droplet-based Enzymatic Assays. Anal Chem *87*, 7938-7943.

Parker, W.B. (2009). Enzymology of purine and pyrimidine antimetabolites used in the treatment of cancer. Chem Rev *109*, 2880-2893.

Pemovska, T., Kontro, M., Yadav, B., Edgren, H., Eldfors, S., Szwajda, A., Almusa, H., Bespalov, M.M., Ellonen, P., Elonen, E.*, et al.* (2013). Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. Cancer Discov *3*, 1416-1429.

Peyssonnaux, C., and Eychene, A. (2001). The Raf/MEK/ERK pathway: new concepts of activation. Biol Cell *93*, 53-62.

Picelli, S., Bjorklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods *10*, 1096-1098.

Pommier, Y. (2013). Drugging topoisomerases: lessons and challenges. ACS Chem Biol *8*, 82-95.

Pon, J.R., and Marra, M.A. (2015). Driver and passenger mutations in cancer. Annu Rev Pathol *10*, 25-50.

Priest, C., Herminghaus, S., and Seemann, R. (2006). Controlled electrocoalescence in microfluidics: Targeting a single lamella. Appl Phys Lett *89*.

Ramensky, V., Bork, P., and Sunyaev, S. (2002). Human non-synonymous SNPs: server and survey. Nucleic Acids Res *30*, 3894-3900.

Rane, T.D., Zec, H.C., and Wang, T.H. (2015). A barcode-free combinatorial screening platform for matrix metalloproteinase screening. Anal Chem *87*, 1950-1956.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47.

Roberts, C.C., Rao, R.R., Loewenberg, M., Brooks, C.F., Galambos, P., Grillet, A.M., and Nemer, M.B. (2012). Comparison of monodisperse droplet generation in flow-focusing devices with hydrophilic and hydrophobic surfaces. Lab Chip *12*, 1540-1547.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rodriguez-Revenga, L., Mila, M., Rosenberg, C., Lamb, A., and Lee, C. (2007). Structural variation in the human genome: the impact of copy number variants on clinical diagnosis. Genet Med *9*, 600-606.

Rowinsky, E.K., and Donehower, R.C. (1991). The clinical pharmacology and use of antimicrotubule agents in cancer chemotherapeutics. Pharmacol Ther *52*, 35-84.

Sackmann, E.K., Fulton, A.L., and Beebe, D.J. (2014). The present and future role of microfluidics in biomedical research. Nature *507*, 181-189.

Sart, S., Tomasi, R.F., Amselem, G., and Baroud, C.N. (2017). Multiscale cytometry and regulation of 3D cell cultures on a chip. Nat Commun *8*, 469.

Schubert, M., Klinger, B., Klunemann, M., Sieber, A., Uhlitz, F., Sauer, S., Garnett, M.J., Bluthgen, N., and Saez-Rodriguez, J. (2018). Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat Commun *9*, 20.

Sequist, L.V., Waltman, B.A., Dias-Santagata, D., Digumarthy, S., Turke, A.B., Fidias, P., Bergethon, K., Shaw, A.T., Gettinger, S., Cosper, A.K.*, et al.* (2011). Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med *3*, 75ra26.

Shembekar, N., Hu, H., Eustace, D., and Merten, C.A. (2018). Single-Cell Droplet Microfluidic Screening for Antibodies Specifically Binding to Target Cells. Cell Rep *22*, 2206-2215.

Sirota, M., Dudley, J.T., Kim, J., Chiang, A.P., Morgan, A.A., Sweet-Cordero, A., Sage, J., and Butte, A.J. (2011). Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med *3*, 96ra77.

Sivagnanam, V., and Gijs, M.A. (2013). Exploring living multicellular organisms, organs, and tissues using microfluidic systems. Chem Rev *113*, 3214-3247.

Solvas, X.C.I., Niu, X.Z., Leeper, K., Cho, S., Chang, S.I., Edel, J.B., and deMello, A.J. (2011). Fluorescence detection methods for microfluidic droplet platforms. Jove-J Vis Exp.

Sonnen, K.F., Lauschke, V.M., Uraji, J., Falk, H.J., Petersen, Y., Funk, M.C., Beaupeux, M., Francois, P., Merten, C.A., and Aulehla, A. (2018). Modulation of Phase Shift between Wnt and Notch Signaling Oscillations Controls Mesoderm Segmentation. Cell *172*, 1079-1090 e1012.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat Methods *14*, 865-868.

Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K*., et al.* (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell *171*, 1437-1452 e1417.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S*., et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Sun, X., Vilar, S., and Tatonetti, N.P. (2013). High-throughput methods for combinatorial drug discovery. Sci Transl Med *5*, 205rv201.

Szalai, B., Subramanian, V., and Saez-Rodriguez, J. (2018). Signatures of cell death and proliferation in perturbation transcriptomics data - from confounding factor to effective prediction. bioRxiv.

Tan, X., Hu, L., Luquette, L.J., 3rd, Gao, G., Liu, Y., Qu, H., Xi, R., Lu, Z.J., Park, P.J., and Elledge, S.J. (2012). Systematic identification of synergistic drug pairs targeting HIV. Nat Biotechnol *30*, 1125-1130.

Thorsen, T., Roberts, R.W., Arnold, F.H., and Quake, S.R. (2001). Dynamic pattern formation in a vesicle-generating microfluidic device. Phys Rev Lett *86*, 4163-4166.

Torre, E., Dueck, H., Shaffer, S., Gospocic, J., Gupte, R., Bonasio, R., Kim, J., Murray, J., and Raj, A. (2018). Rare Cell Detection by Single-Cell RNA Sequencing as Guided by Single-Molecule RNA FISH. Cell Syst *6*, 171-179 e175.

Turke, A.B., Zejnullahu, K., Wu, Y.L., Song, Y., Dias-Santagata, D., Lifshits, E., Toschi, L., Rogers, A., Mok, T., Sequist, L*., et al.* (2010). Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell *17*, 77-88.

Umbanhowar, P.B., Prasad, V., and Weitz, D.A. (2000). Monodisperse emulsion generation via drop break off in a coflowing stream. Langmuir *16*, 347-351.

Unger, M.A., Chou, H.P., Thorsen, T., Scherer, A., and Quake, S.R. (2000). Monolithic microfabricated valves and pumps by multilayer soft lithography. Science *288*, 113-116.

Utharala, R., Tseng, Q., Furlong, E.E.M., and Merten, C.A. (2018). A Versatile, Low-Cost, Multiway Microfluidic Sorter for Droplets, Cells, and Embryos. Anal Chem *90*, 5982-5988.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. Science *339*, 1546-1558.

Vyawahare, S., Griffiths, A.D., and Merten, C.A. (2010). Miniaturization and parallelization of biological and chemical assays in microfluidic devices. Chem Biol *17*, 1052-1065.

Wang, J., Fan, H.C., Behr, B., and Quake, S.R. (2012). Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. Cell *150*, 402-412.

Wei, G., Twomey, D., Lamb, J., Schlis, K., Agarwal, J., Stam, R.W., Opferman, J.T., Sallan, S.E., den Boer, M.L., Pieters, R.*, et al.* (2006). Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. Cancer Cell *10*, 331-342.

Weinberg, R.A. (1991). Tumor suppressor genes. Science *254*, 1138-1146.

Whitesides, G.M. (2006). The origins and the future of microfluidics. Nature *442*, 368-373.

Woo, J.H., Shimoni, Y., Yang, W.S., Subramaniam, P., Iyer, A., Nicoletti, P., Rodriguez Martinez, M., Lopez, G., Mattioli, M., Realubit, R.*, et al.* (2015). Elucidating Compound Mechanism of Action by Network Perturbation Analysis. Cell *162*, 441-451.

Ye, C., Ho, D.J., Neri, M., Yang, C., Kulkarni, T., Randhawa, R., Henault, M., Mostacci, N., Farmer, P., Renner, S.*, et al.* (2018). DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery. Nat Commun *9*, 4307.

Yu, L., Chen, M.C., and Cheung, K.C. (2010). Droplet-based microfluidic system for multicellular tumor spheroid formation and anticancer drug testing. Lab Chip *10*, 2424-2432.

Zhang, J., Adrian, F.J., Jahnke, W., Cowan-Jacob, S.W., Li, A.G., Iacob, R.E., Sim, T., Powers, J., Dierks, C., Sun, F.*, et al.* (2010). Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors. Nature *463*, 501-506.

Zhang, X., Li, T., Liu, F., Chen, Y., Yao, J., Li, Z., Huang, Y., and Wang, J. (2018). Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. bioRxiv.

Zorita, E., Cusco, P., and Filion, G.J. (2015). Starcode: sequence clustering based on all-pairs search. Bioinformatics *31*, 1913-1919.

9. Appendix

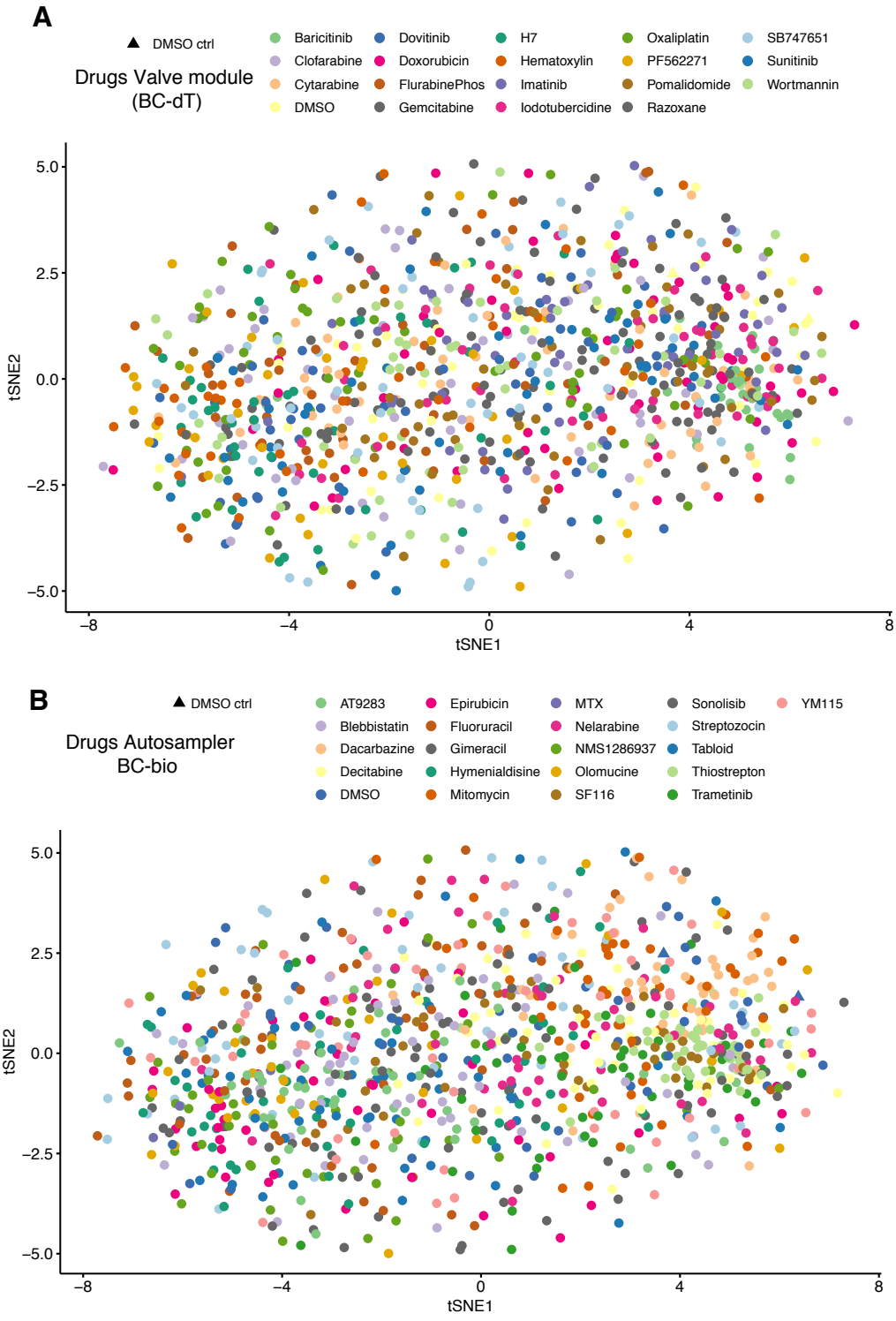

Fig 9.1: Cluster analysis screen 1 with improved library preparation

(A) Cluster analysis using tSNE of gene expression data from 399 treatment conditions colour coded according to drugs from the valve-module encoded by BC-dT. **(B)** Cluster analysis of gene expression data from (A) colored according to the 21 drugs and DMSO injected by the autosampler encoded by BC-bio

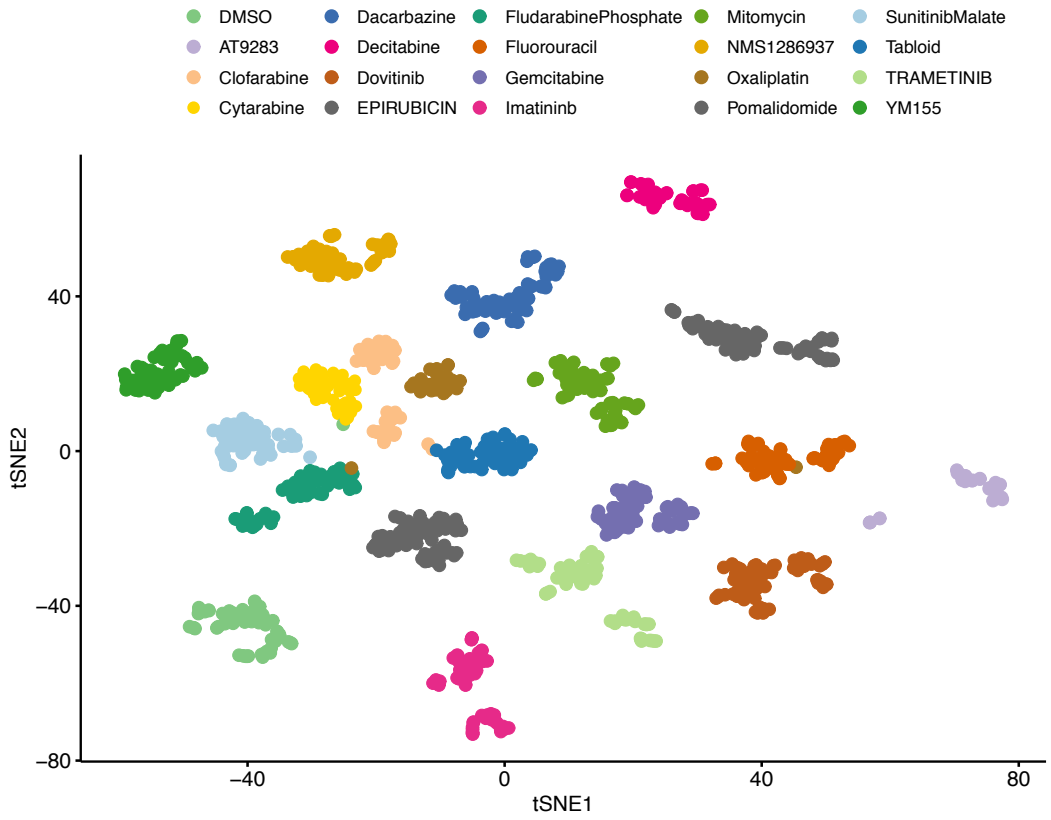

Fig. 9.2: Cluster analysis after 6h drug exposure

tSNE based analysis of gene expression data of cell treated with drug combination for 6h. Each data point corresponds to one treatment condition. Data points are coloured according to drug injected from the valve-module which were encoded by BC-dT

10. Abbreviations

AS	Autosampler
BC	Barcode
BC-bio	Biotinylated Barcode
BC-dT	Poly-dT Barcode
Bio	Biotin
C	Celsius
CV	Coefficient of Variation
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
dpi	Dots per inch
dT	Desoxythymidin
EGFR	Epidermal Growth Factor Receptor
FADS	Fluorescence activated droplet sorting
FBS	Fetal bovine serum
Fwd	Forward
GR	Growth reduction
h	Hours
min	Minutes
mm	Millimetre
mRNA	messenger Ribonucleic Acid
ms	Milliseconds
NGS	Next Generation Sequencing
NSCLC	Non-small cell lung cancer
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PDMS	Poly-Di-Methyl-Siloxane
PFO	1H,1H,2H,2H-Perfluorooctanol
PMT	Photomultiplier Tube
PMT	Photomultiplier tube

PS-1	Pico-Surf 1
PTFE	Polytetrafluorethylen
qPCR	quantitative Polymerase Chain Reaction
Rev	Reverse
RNA	Ribonucleic Acid
RNA-Seq	RNA sequencing
rpm	Revolutions per minute
RT	Reverse Transcription
s	Seconds
SDS	Sodium dodecyl sulfate
sec	Seconds
tSNE	t-distributed Stochastic Neighbourhood Embedding
UV	Ultraviolet