Genetic interactions in pediatric cancer – Current state and outlook

Final version 7-9-2021; Richard Gremmen (student no 6642896); examiner: Josephine Daub, 2nd reviewer: Roland Kuiper

Layman's Summary

Kanker is een ziekte van het DNA. Mutaties in het DNA zorgen ervoor dat genen zich anders gaan gedragen. Ze kunnen er aan bijdragen bij dat cellen zich als kankercellen ontwikkelen, bijvoorbeeld door ongebreidelde groei te vertonen. Genen opereren echter niet altijd onafhankelijk, ze hebben invloed op elkaar. Wanneer dat zo is, spreken we van genetische interacties. Dat leidt ertoe dat sommige mutaties vaak samen optreden omdat ze een voordeel voor de tumor opleveren, of dat mutaties juist minder vaak samen voorkomen (of helemaal niet) omdat de combinatie van mutaties ziekmakend of dodelijk is voor de kankercel. De laatste situatie noemen we synthetische dodelijkheid. Het legt een kwetsbaarheid bloot van een tumor die al één van de mutaties heeft. Door nu op te zoek te gaan naar deze combinaties kan de ontwikkeling van kanker beter begrepen worden en kunnen er mogelijke nieuwe behandelmethodes ontwikkeld worden. Voor volwassen kanker is dat al gelukt met gerichte therapie met een medicijn dat als PARP-remmer werkt (PARP is een gen). Daar kunnen patienten met een tumor met een mutatie in een ander gen, het BRCA1- of BRCA2-gen baat bij hebben. Voor kinderkanker is zoiets nog minder toegepast. Omdat er grote verschillen tussen kinder- en volwassenkanker zijn kunnen behandelingen ook niet makkelijk overgenomen worden. Dit verslag gaat over de zoektocht naar nieuwe genetische interacties die specifiek voor kinderkanker zijn. We bespreken de experimentele en statistische methoden en de resultaten die daarmee behaald zijn. We zien dat kinderkanker zeldzamer is en er daardoor ook minder data beschikbaar zijn die in de zoektocht kunnen helpen. Het kan helpen om ook andere soorten mutaties in het DNA en afwijkend gedrag van genen mee te nemen, niet alleen de kleine mutaties van het DNA, om meer data beschikbaar te hebben. Het kan ook voordeel leveren data van meerdere onderzoeksinstituten te combineren. In het verslag worden een aantal mogelijkheden en uitdagingen genoemd om deze data optimaal te combineren, zodat de zoektocht naar genetische interacties in kinderkanker bespoedigd wordt. We verwachten dat studies die hier gebruik van maken nieuwe genetische interacties in kinderkanker zullen vinden en uiteindelijk kunnen helpen bij de ontwikkeling van nieuwe behandelingen.

Abstract

Cancer is a genomic disease where single driver mutations and combinations of mutations determine cancer progression. In this review we discuss the search for genetic interactions in pediatric (childhood and young adult) cancers. Genetic interactions imply that the contributions of (mutations in) genes to phenotypes are not independent and influence each other. This allows for detection of vulnerabilities of malignant cells relevant to clinical interventions, especially if combinations of two mutations are lethal to the cell (synthetic lethality). Such vulnerabilities have been found for adult cancers and have been translated to clinical applications in some cases. Research has revealed many molecular causes of cancer progression and genetic interactions between genes and between pathways. However, there are large differences between adult and pediatric cancers in these causes. For example, where adult cancers mostly slowly accumulate cancer driver mutations, pediatric cancers seem to arise from a specific mutation at a certain time during development. Consequently, therapeutics based on these genetic interactions for adult cancers are not fit for pediatric cancer types having their own specific interactions. Detecting new interactions will help to understand the progression and treatment of pediatric cancer. There are both experimental and computational methods to find candidates for interacting genes. The computational detection of these gene pairs depends on whether you search for them in a dataset with multiple cancer types or in a subset with only samples from one cancer type. Furthermore, the rare occurrence of pediatric cancer compared to adult cancers is challenging, because statistical tests for the detection of genetic interactions have less power. Therefore it is probable that many clinically relevant interactions have yet to be discovered. There are multiple ways to increase statistical power: first, we can combine data from multiple clinics. Second, existing analyses mostly take only small mutations in the genome into account, so adding other kinds of mutations may improve detection. There are many ways to combine these data, either by adding samples, enriching sample data with other kinds of mutations or combining the outcome of predictive models based on separate data sets. The newly discovered interacting gene pairs need to be validated in targeted experiments under controlled conditions. Studies, following the improvements discussed in this review, will find many more genetic interactions in pediatric cancer, likely resulting in new therapeutic targets.

Introduction

Cancer is a genetic disease. Much progress has been made in treating certain cancers by targeting single driver genes or using immunotherapies. Numerous genetically targeted cancer therapies have been developed but they are restricted to a relatively small set of cancer driver genes. Some targeted cancer therapies exploit specific weaknesses of cancer cells (Brody 2005, McLornan 2014) based on the search for lethal combinations of existing alterations, such as the successful application of PAPR inhibitors in BRCA1/2-deficient patients (Bryant 2005, Farmer 2005). Relatively recent genome editing techniques like CRISPR-Cas9 and ever cheaper and more effective sequencing techniques enable the search for other combinations of mutations that can specifically kill cancers cells while leaving healthy cells unharmed (review Huang 2020). Such “synthetic lethal” gene pairs belong to a wider class of genetic interactions (GI) between genes. They may result in new treatments for *pediatric* cancers. Three in four children with pediatric cancer survive by successful

treatments, but many of those have negative effects in later stages of life. The central questions of this review are: what genetic interactions have been found in pediatric cancers, what makes them specific, how many more are expected to be found and how can we accelerate this search? First, we recapitulate definitions and terms of genetic interactions. We describe experimental and computational methods used for detecting these genetic interactions in general and their results so far in adult cancers. We discuss how features of pediatric cancers differ from those of adult cancers and impact detection of genetic interactions. Next, we assess which pediatric data sets are available with what kind of data. Also, we describe the current results for genetic interactions in pediatric cancers based on the aforementioned methods in these data sets. Finally, we give an outlook and reflect on possible new ways and challenges for detecting genetic interactions in pediatric cancer.

Definitions in genetic interactions

Positive and negative interactions

Genetic alterations in cancer can have greater or weaker consequences when combined with the effects of alterations in other genes. Definitions of interaction are always relative to a particular quantitative phenotype, for example the expression of a certain protein, the amount of a metabolic product, the growth rate or another proxy value for fitness. Mani et al reviewed formal definitions: the multiplicative (or log-additive model) is useful and predicts that double-mutant fitness is the product of the single-mutant fitness values (Mani 2008; figure 1a). Based on this model, unexpected outcome of genetic mutations on genetic background is known as epistasis or genetic interaction (Lehner 2011; Domingo 2019). Such an unexpected outcome resulting from mutations in just two different genes is a gene-gene interaction and is the focus of this review. Interactions can be positive or negative. In case of gene-gene interaction the resulting phenotype of the two mutations in the genes deviates from a log-additive model and is nonlinear and non-additive (Kafaie 2020, Domingo 2019). We consider the gene-gene interaction to be positive or negative if the effect on the phenotype is respectively larger or smaller than expected based on the effects of the single genes on the considered phenotype (figure 1a, b). In this review the terms “genetic interaction between genes”, “gene-gene interaction” and “gene pairs” always refers to the interaction of any mutations or alterations in two genes. Sometimes, positive interactions are referred to as aggravating and negative interactions as alleviating (Boettcher 2018). All these terms only inform about the quantitative direction of the effect of the interaction in a phenotypical quantity, and not per se on the more general effect on fitness of the cell (reflected by ecological statistics). We are mainly interested in the latter with potential clinical applications for treatment and prognosis in mind. The strength and direction of the interaction of two mutations is an important factor. Every outcome that deviates from an expected outcome is considered to be a genetic interaction: it can be only in magnitude, it can be masking (masking epistasis), the sign can reverse and it may even reverse the effect of both mutations (reversal/reciprocal sign) (figure 1c).

Functional events and underlying causes

What type of alterations do we consider? The phenotypic effect of a single mutation can be caused by gene loss, gain of function, a dosage effect or any other molecular mechanism. The molecular mechanisms underlying these quantitative effects start with either a genetic or

an epigenetic alteration. Genetic alterations in sequences can be single nucleotide variants (SNVs) or events caused by structural variants (SV) like gene fusions, gene duplications or copy number variations (CNV), inversions and translocations. These may also result in different regulation by non-coding sequences. Epigenetic alterations (i.e. without alterations in sequences) like methylation may also result in changes in gene expression. All of these events can result in a changing phenotype, for example gene silencing by methylation can result in gene loss and copy number variations can result in dosage effects. We include all of these events in this review and we use the words mutation and alteration (in this general sense) interchangeably.

Mutual exclusive and co-occurring pairs as possible indication for interactions

Statistical evidence can only point to the molecular mechanisms of interactions which ultimately result in an unexpected phenotypic value. Therefore, it is important to discriminate between terms for these *mechanisms* and terms that refer to *detection* of interactions in genome-wide experimental or computational screens (to be discussed later). Researchers aim to find significant statistical evidence for the interactions. Some mutations often occur together in the same malignant cell or (heterogeneous) sample or tumor. We refer to these mutated gene pairs as co-occurring (CO), if they occur more frequent together in gene pairs than expected by chance alone. Such a finding can be an indication of a positive interaction. Alternatively some mutations seldom occur together, and we refer to them as mutually exclusive (ME) mutations if occurrence happens less often than expected by chance (Lehner 2011), pointing at a putative negative interaction.

Mutual exclusivity is a possible indicator of a negative interaction, but may also be caused by coincidence. Different cancer subtypes each having their own biological processes may also result in detection of ME pairs. Additionally, there may be a detection bias by experimental methods or other technical reasons resulting in ME pairs without a genetic interaction as their cause. Similarly, co-occurring gene pairs can suggest positive interactions among other possible causes. To summarize, a detected statistical irregularity like ME and CO can have various reasons with a genetic interaction as only one possible cause among others. If the genetic interaction is the main contributor, the exact molecular mechanisms underlying the interaction have still to be revealed to connect those to the statistics.

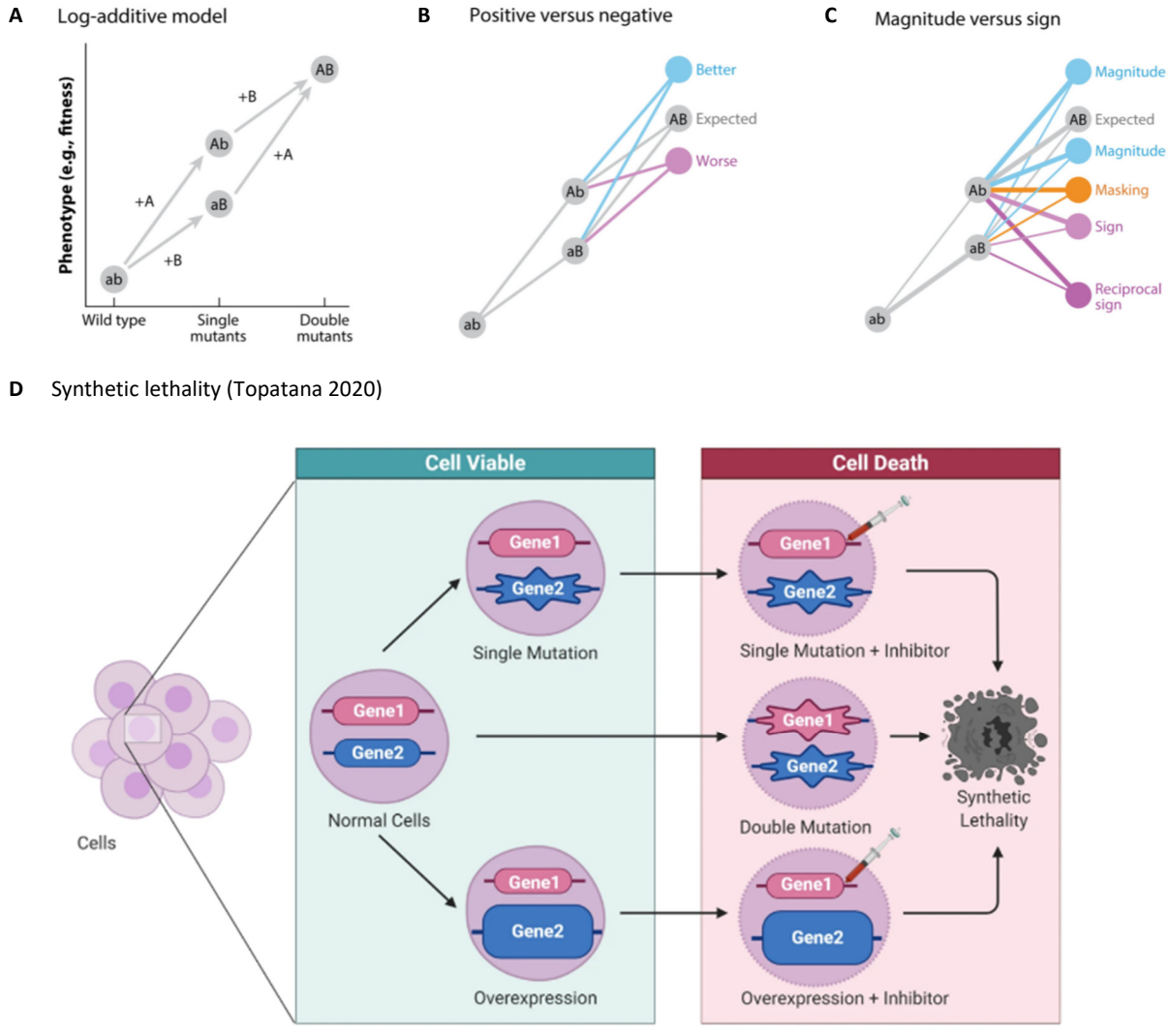

Figure 1: Strength and direction of genetic interactions and synthetic lethality

A Log-additive (or multiplicative) **model** for independent (non-interacting) mutations in two genes: multiplicative mean of impacts of the A and B alleles. Double-mutant fitness is the product of the single-mutant fitness values.

B Outcomes of two mutations that deviate from the expectation as shown in A. **Positive** or **negative interactions** show a larger or smaller contribution respectively to the phenotype than expected based on the log-additive model in **A**.

C magnitude epistasis: the genetic interaction results in a change in magnitude of the effect of a mutation, but the direction stays the same. **Masking epistasis** is defined as complete indifference of the effect of one mutation to the magnitude. **Sign epistasis** changes the direction of the effect. **Reciprocal sign epistasis** occurs when this happens for both alterations.

(**A-C** adapted from Domingo 2019)

D Adapted from Topatana 2020. **Synthetic lethality** can occur due to various effects of a single alteration, for example complete gene loss and zero expression, gain-of-function resulting in overexpression, as long as the combined alterations result in non-viable cells (zero fitness).

Molecular causes of genetic interactions

Possible molecular causes and mechanisms of epistatic interactions have been extensively described (reviewed in Lehner 2011 and Domingo 2019). Examples are alterations resulting in a different binding affinity to another molecule, a change in protein or protein cluster configuration etc. As discussed, these molecular changes may result in quantitative phenotypic changes. The exact molecular mechanisms underlying mutual exclusivity can be divided in two general causal groups. First, it may be caused by a selective disadvantage of having both alterations, a "synthetic sickness" (Ashworth 2011). An extreme case of such a

selective disadvantage is synthetic lethality, that means that cells having both alterations are not viable but having one of the alterations has no negative effect (figure 1d). Second, the genes may have redundant functions or redundantly alter the same pathway (*intra*-pathway epistasis) or are part of two different redundant pathways (*inter*-pathway epistasis), and mutation of both genes as opposed to mutation of only one gene has no further selective advantage. Apart from the aforementioned general caveats in explanation of ME pairs, it is difficult to distinguish between these two specific biological explanations (O'Neill 2017).

Co-occurrence can be caused by a selective advantage, for example by cooperative interactions between oncogenes. Both mutations may either be neutral/beneficial, but the combination gives an (even higher) advantage. This scenario is the most likely. Rarer is when one or both mutations has a negative effect, but one mutation compensates the negative effect of the other mutation (compensation): it masks the negative effect or completely rescues the cell (synthetic viability). One of the possible mechanisms is that the second mutation keeps an effective binding at the protein-protein interface (conformational epistasis; Ashworth 2011).

From genetic interactions to therapeutic targets

Knowledge of genetic interactions and the underlying mechanisms of the interaction can contribute to a fundamental understanding of the development of cancer and pathways involved in the development. The search for co-occurring mutations and modules of drivers gives insight into cooperation of genes and pathways. Genetic interactions in general can reveal different subtypes of cancer (as we will explain later). These subtypes are clinically relevant, because each subtype may need another treatment. The main goal of scanning for mutual exclusive gene pairs is to detect vulnerabilities for malignant cells. ME pairs with the underlying cause of synthetic lethality (figure 1d) may be exploited as a potential therapeutic target by inhibiting one of the involved genes in patients that have a mutation in the other gene.

Methods

The most direct and unsupervised method to find genetic interactions is to use genome-wide experimental perturbation screens. However, this also has limitations in translation to the clinic. The next closest (computational) method is to find structure in the statistics of existing mutations and expression values in cancer samples. These computational methods aim to find mutual exclusive and co-occurring pairs as possible indication for interactions. We discuss both and focus on the latter computational methods.

Experimental perturbation methods: CRISPR/RNAi screens and validation experiments

Genome-wide perturbation experiments in cell lines can experimentally determine genetic interactions and synthetic lethality. One introduces double knock-outs and looks at the effects on phenotypes like cell growth, cell viability or expression in comparison to the effects in single mutants. RNA interference (RNAi) was used in the earliest screens, and currently CRISPR-Cas editing has predominantly been used (review Dhanjal 2017; O'Neil 2017). For example, automatic interaction screens using RNAi were applied to study the interaction of epigenetic regulators with a set of genes in human colon cancer cells by measuring the effect on multiple phenotypic quantities (Laufer 2013). Dual gene inactivation (knockout) screens

using a dual-Cas9 (CRISPR-Cas9) system were able to reveal synthetic lethality pairs in various human cancer cell lines (Najm 2018). A comparable method revealed directionality of genetic interactions in human K562 CML leukemia cells (Boettcher 2018). Some methods combine multiple kinds of data, like the SELECT framework that filters gene pairs based on data from CRISPR screens, patient prognosis data, phylogenetic knowledge (known/suspected interactions in homologues of other species) and drug response data (Lee 2021).

Next to these genome-wide screens, one can also perform focused synthetic lethality experiments in vitro (in patient-derived cell cultures or organoids and cell lines) to validate gene pair candidates earlier identified in computational methods. This may also help to investigate potential molecular mechanisms, e.g. for the genetic interaction of the TP53/DROSCHA-pair in Wilms tumor (Daub 2020, and Verhagen et al, private correspondence).

Exploiting existing mutations: computational methods

Deng et al reviewed several in silico methods and discriminated between *de novo* methods using only genomic data and knowledge-based methods that additionally incorporate extra information, such as signaling and regulatory pathways, interaction networks or phenotypes (Deng 2019).

All *de novo* methods aim to detect if some pairs of genes are more or less often mutated together than by chance, reasoning that there is a selective pressure working on the alterations. Kafaie et al reviewed methods to identify combinations of single SNVs. This is a highly dimensional problem on sparse data with high computational costs (Kafaie 2020). Therefore, it makes sense to turn to less granular methods at the (coarser) gene level with a gene marked as altered when *any* functional event in that gene occurs, resulting in less sparse data (Park 2015, Kim 2017, and Daub 2020). The simplest method is Fisher's exact test, which is a crude approximation because it assumes an independent and identical distribution of mutations, i.e. it does not take varying probabilities across samples and genes into account (Canisius 2016). Another category of methods on gene level consists of permutation methods on binary matrices consisting of samples as rows and genes as columns. Each element in the matrix either contains a one or a zero based on whether or not there is an alteration in the gene in a sample. Note that these gene/sample matrices are irrespective of a putative underlying GI caused by a mutation, and its direction or strength, and there may be (noisy) averaging out of the effects of various mutations in the same gene. Subsequently these matrices are permutated: the binary gene information is reassigned to random samples in the data set while keeping the number of counts fixed per row and column and thus taking into account different alteration frequencies across samples and genes. Subsequently one can determine statistical significance of the pairwise association between genes in the original matrix as compared to the permutated matrix. Permutation methods can be applied irrespective of the type of genetic alteration, so when evaluating results of studies applying these methods, it is important to take into account what types of events were included in the study. A permutation method creates a null distribution of mutations across genes and samples and its performance depends on the number of genes and samples included and can be very time-consuming.

Various computational methods have been published. In a method called "Weighted Sampling based Mutual Exclusivity" (WeSME) Kim et al improved on computational costs by using weighted sampling based on mutation frequency in each sample to detect candidates for negative interactions, enabling to include more genes in the analysis (Kim 2017). Leiserson et al used another method for weighted sampling (WExT) that is more computationally expensive but may also be able to test in hyper-mutated samples (Leiserson 2016). Some other methods are parametric, as they assume a certain distribution of mutations. One example is the method DISCOVER ("Discrete Independence Statistic Controlling for Observations with Varying Event Rates") that simulates a theoretical distribution and is computationally cheaper (Canisius 2016). Variations of the parameters allow for various distributions of single alterations in genes over samples. Canisius et al showed that DISCOVER better takes heterogeneous alteration rates across samples into account than simple methods like Fisher's exact test and could avoid the negative effects of those simple methods resulting in a high false discovery rate for co-occurring pairs and an underestimation of ME pairs (Canisius 2016).

Alterations in single genes with high mutation rates pass detection threshold more easily than low-mutating genes (O'Neil 2017). Therefore false negatives are sooner to be expected for gene pairs containing low-mutating genes especially in small data sets resulting in tests with relatively small statistical power.

Aggregating towards (unsupervised) gene modules (or subnetworks) can further increase statistical power because alterations of single genes in a gene module are taken together. Inferring de novo mutual exclusive (ME) modules depends on graph-based methods. They aim to detect cliques in bipartite gene-sample networks and are computationally expensive. This has been partially resolved with heuristic methods, for example by Vandin et al (Vandin 2011) and in a method called HotNet2 where the contribution of a mutation in a single gene is dispersed over module neighbors (Leiserson 2015). Kim et al developed a different method MEMCover which could also identify new ME modules playing a role in cancer types (Kim 2015). ME modules can point to pathway epistasis, i.e. either one pathway or the other has been affected: the (alterations in) pathways could be either redundant or synthetically lethal.

Computational methods can also study correlations to other clinical data than mutational occurrences alone. Some co-occurring events or expression levels are predictive for either poor or favorable outcome in patient survival (review in Sweet-Cordero 2019) and these correlations may contain a signal for prediction of gene pair candidates. An example is the recurrent structural *alterations* in acute myeloid leukemia (AML) related to survival (Bolouri 2018). Magen et al correlated pairwise relationships in gene *expression* to patient survival (Magen 2019).

Knowledge-based methods

Knowledge-based approaches that integrate *known* pathways can alleviate computational costs and have the additional advantage of being more interpretable (reviewed in Deng 2019; methods evaluated in Reyna 2020). They can reveal *inter*-pathway epistasis of known pathways. However, they are not able to find (interactions between) unknown gene modules. Tests on interactions between pathways may have a higher statistical power and may be able

to find rare interactions not detectable by analysis on gene level. This also applies to the aforementioned unsupervised gene modules. However, gene pairs *within* a gene module or pathway (*intra*-pathway epistasis) will no longer be found, because sensitivity on gene level will be lost. Also, incorporating external information may be a confounder when not specific enough. For example, Deng et al stressed that one needs tissue-specific instead of general networks, if one aims to detect tissue-specific functional effects of gene interaction. This is caused by wildly varying expression profiles among tissues and cell types. Other possible additions are integration of gene co-expression, functional profiles and drug response profiles (e.g. method REVEALER, J.W. Kim 2016).

Results in adult cancers

Various synthetic lethal pairs have been found in adult cancers including oncogene drivers (e.g. BRAF or KRAS) and non-oncogene "addictions", where a mutation in an oncogene depends on another gene for cell survival (Kandoth 2013, McLornan 2014, Park 2015). The ultimate example of successful use of synthetic lethality is the development of PARP inhibitors (reviewed in Underhill 2011) based on disturbance of DNA repair pathways. Base-excision repair (BER) of single-stranded breaks (SSB) occurs often in cells. It is known since 1995 (Lindahl 1995) that PARP1 plays a key role by detecting and binding to SSBs and recruiting multiple other proteins needed for repair. PARP inhibitors result in cells with a lot of SSBs that are converted into double stranded breaks (DSBs) during replication. Bryant et al and Farmer et al found that the use of PARP inhibitors could kill BRCA1/2 deficient tumor cells (Bryant 2005, Farmer 2005). Only cells having both inhibition of PARP and mutations in BRCA die: BRCA-deficient cells have a malfunctioning homologous recombination (HR) pathway and this error-free pathway can no longer rescue the errors originally invoked by the PARP inhibitors. Alternatively used and less error-free pathways accumulate errors that ultimately result in cell death. This success inspired the search for other synthetic lethal pairs as possible clinical targets. Much focus has been on DNA damage repair (DDR) pathways (McLornan 2014). Topatana et al give an overview of the status of clinical trials exploiting many SL gene pairs in DDR pathways (Topatana 2020).

Some more (non-extensive) examples other than in DDR pathways give an idea of the extent of interactions found so far with various computational and experimental methods. Park et al studied data of 22 different (not specifically adult or pediatric) tumor types from The Cancer Genome Atlas (TCGA) Research Network (Ciriello 2013) and found 30 positive interactions and 25 negative interactions in a pan-cancer analysis, and 60 interactions in an analysis within a single cancer type (Park 2015). Deng et al (Deng 2019) named a few examples of ME pairs that yielded a favorable outcome in clinical trials, for example mutations in KRAS and EGFR for lung adenocarcinomas (Unni 2015), WT1 and TET2 in acute myeloid leukemia (AML) (Wang 2015), and BRAF and NRAS in melanoma (Davies 2002) among several other adult cancers. Najm et al found new gene pairs in experimental double knock-out screens consisting of MAPK pathway genes, apoptotic genes and TP53 (Najm 2018). Cui et al found interaction partners of EFGR based on survival analysis in experimental screens (Cui 2021). Interestingly, Park et al found that more than 90% of the genetic interactions were only detected in a single cancer type (Park 2015). Leiserson et al showed that interactions between *de novo* detected gene modules were also mostly cancer type specific (Leiserson 2015).

Therefore it is plausible that pediatric cancers also contain specific gene pairs and pathway interactions that differ from those in adult cancers.

Pediatric data sets

As mentioned in the introduction, there are major differences between pediatric and adult cancers. Before discussing these differences, we introduce some pediatric data sets confirming these findings. We focus on genomic data about alterations in pediatric cancer, such as Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES) and on RNA expression (RNA-seq) data. WGS has the advantage of finding mutations in regulatory non-coding sequences (e.g. mutations in transcriptional factors) as compared to WES. WGS can also detect large chromosomal rearrangements. RNA-seq data can be used to confirm alterations in the WGS or WES data and gene fusion products. Additionally, expression data reveal if a gene is actually expressed and this includes epigenetic effects like gene silencing by methylation or suppression and activation by regulatory networks. Apart from these, there are data sets about phenotypic data, e.g. drug response data, patient outcome data etc. that we do not include in our review.

Most pediatric cancer studies are cancer type specific, for example for neuroblastomas (Valentijn 2015). Efforts to collect pediatric cancer genomes started with Pediatric Cancer Genome Project (PCGP) of St. Jude and Washington University (Downing 2012). The first two large genomic data sets published for pediatric cancer were the DKFZ and TARGET data sets. One is from the German Cancer Consortium (DFKZ) containing whole exome (WES) or whole genome (WGS) data for 24 types of (mainly central nervous system (CNS) tumors (Gröbner 2018). The TARGET data set from the National Cancer Institute (NCI) for four blood and two solid pediatric cancers contains WGS or WES data and some RNA expression sequencing data for B-lineage acute lymphoblastic leukaemia (B-ALL), T-lineage acute lymphoblastic leukemia (T-ALL), acute myeloid leukaemias (AML), neuroblastomas (NBL), Wilms tumor and osteosarcoma (Ma 2018). Both data sets complement each other, because the DFKZ data set contains relatively few blood cancer samples for the cancer types in TARGET.

In Australia pediatric cancer samples are collected in the Zero Childhood Cancer Program (ZERO) https://www.zerochildhoodcancer.org.au. There are also specialized collaborations like The Children's Brain Tumor Tissue Consortium (CBTTC) (https://cbttc.org).

Differences between pediatric and adult cancers

In this section we discuss the differences between adult and pediatric cancers. Sweet-Cordero et al have given a more complete overview (Sweet-Cordero 2019). These are relevant in respect to the exact data to be included, the best methods of statistical modelling and computing to detect specific genetic interactions, the correct analysis of the results and the challenges to expect.

Mutational load: number of driver and passenger mutations

Most genes in (surviving) cancer cells are affected by point mutations (Vogelstein 2013). It is well established that most pediatric cancers have less single nucleotide variants (SNV), in general mainly believed to be passenger mutations. Ma et al found 0.17-0.79 mutations/Mb in pediatric cancers vs 1-10 in adult cancers (Ma 2018). Gröbner et al found at least one

significantly mutated gene (SMG, pointing to a driver gene) in 47% of pediatric (mainly CNS) cancers, while in contrast at least one SMG was found in 93% of adult cancers (Gröbner 2018). In most adult cancers, an average of two to eight sequential alterations are necessary for full development until metastasis (Vogelstein 2013). A classic example is the multi-hit model of the progression of colon cancer, where four alterations in the driver genes APC, TP53, KRAS and SMAD4 result in full progression (Fearon 2011). This has been studied extensively and could be reproduced in vitro (Drost 2015). However, in pediatric cancers sometimes only one or two driver mutations are identified (Vogelstein 2013).

Pediatric cancer as a developmental disease

How do pediatric cancers develop despite the lesser number of SNVs? First, many childhood cancers can be considered to be developmental diseases , where cells take a wrong turn in their development and stay in a state resembling that of stem cells. These events are believed to be conditioned on the developmental stage and the cells harboring them are for example already prone to continuous proliferation, and therefore need less additional driver gene mutations than adult cancers to become malignant. Mutations in developmental pathways such as WNT signaling occur often in solid pediatric tumors (Parsons 2016). It is hypothesized that due to germline or somatic mutations, progenitor cells in various stages of their developmental trajectories become cancerous (Downing 2012). These progenitor cells already have characteristics that trigger cancer progression, as in cancers of the blood like juvenile myelomonocytic leukemia (JMML) (Solman 2000). Also, pediatric cancers often arise from (or during) developing mesodermal tissue during early organ formation, growth and maturation (Downing 2012).

The recent revolution in single-cell sequencing has revealed many developmental trajectories in development and cancer progression. Single-cell RNA sequencing has also helped to decipher the various consecutive cell types in developmental lineages. It is now possible to deduce the cell-of-origin among the various progenitor cells. Kim et al found that cell-of-origin features also dominated the expression state of various adult tumor types (Kim 2015). This has been explicitly shown for pediatric cancers: the difference between T-ALL and B-ALL, is based on separate clusters in gene expression resulting from different origins. Other cancer types for which the cell-of-origin could be established are neuroblastoma (Hanemaaijer 2021, Kildisiute 2021), malignant rhabdoid tumours (MRT) (Custers 2021) or JMML (Solman 2020) among others. Kildisiute et al discovered a common transcriptional state in malignant cells of various cell types in the same tumor, indicating that the cancer started in a common cell-of-origin (symphatoblast; Kildisiute 2021).

Epigenetic features

Related to the view of pediatric cancer as a developmental disease is the role of epigenetic alterations promoting cancer development, that are known to happen early in development and happen relatively often in pediatric cancers (reviewed in Filbin 2019). Many pediatric cancers have mutations in genes encoding epigenetic regulators, for example for hematological malignancies. Izzo et al showed that early mutations in Tet2 and DNMT3A in hematopoietic stem and progenitor populations result in disruption of genome-wide methylation, changes in (CpG rich) transcription factor binding and resulting changes to

developmental trajectories of cells (Izzo 2020). There are mainly two ways for cancers to change expression of genes by epigenetic mechanisms: i) modification of proteins that change the spatial conformation of DNA and its interactions with proteins (e.g. histone modification) and ii) modifications of the nucleotides in DNA like methylation among others (Locke 2019). An example of the first class of modifications is the loss of SMARCB1 in malignant rhabdoid tumors (MRT). SMARCB1 is part of SWI/SNF complex involved in chromatin remodeling, is a regulator of chromatin accessibility and acts as a tumor suppressor (Calandrini 2020, Custers 2021). An example of the second class is the methylation of the MGMT (methylation of O6-methylguanine-DNA methyltransferase (MGMT) promoter. This specific methylation has been used as a biomarker, is related to patient's response to drugs (Locke 2019) and is relevant for tumor characterization and treatment: Weidema et al showed for the (adult) angiosarcomas that MGMT hypermethylation can differ between clusters (based on methylation profiling) (Weidema 2020). Capper et al showed that methylation profiles can be used to predict cancer types with high accuracy (Capper 2018). However, DNA methylation may also be heterogeneous within certain cancer types, as was shown in glioblastoma (Wenger 2019).

Germline mutations and hereditary predisposition

It is important to include germline mutations in the search for interactions. Germline mutations play a substantial role in both pediatric and adult cancer. Zhang et al found that approximately 8% of *pediatric* cancer patients have a hereditary predisposition through germline mutations (versus about 1% in healthy adults) in a data set of a selected set of cancer types dominated by leukemia. (Probable) pathogenic germline mutations were mostly found in non-CNS solid tumors and CNS tumors, in that order, and the most occurring mutation was found in the tumor suppressor gene TP53 (Zhang 2015). Parsons et al found a correlation of germline mutations to patient phenotype in 10% of solid pediatric tumor cases and concluded that genetic testing could be improved (Parsons 2016). For this, Byrjalsen et al proposed new gene selection criteria for gene panels (Byrjalsen 2021). Germline mutations contribute to pediatric cancer being a developmental disease, for example by deregulating RAS signaling in JMML (review Niemeyer 2019).

In 33 *adult* cancer types, Huang et al and al found rare germline mutations in 8% of tumors (Huang 2018). Carter et al found multiple interactions between germline and somatic mutations and linked them to biological models, for example between a germline mutation in RBFOX1 and a somatic mutation in SF3B1 that plays a role in splicing (Carter 2017). They also observed pathogenic variants earlier discovered by Zhang et al in pediatric cancers (Zhang 2015).

Specific tumor types, driver mutations and pathways

Driver events can depend on the developmental stage and may therefore be specific for a tissue type developing from a developmental compartment (Sweet-Cordero 2019).

Both for adult and pediatric cancer, cancer driver genes and pathways have been found to be specific for tissues (Kim 2015). In pediatric cancers this is true for "Wilms tumor, T cell acute lymphoblastic leukemia (T-ALL), and acute myeloid leukemia (AML)" among others (reviewed

in Sweet-Cordero 2018). Some tumor types, like brain and solid tumors that develop in children are very rare in adults, e.g. Wilms tumor.

Driver mutations have often been found to be specific for pediatric cancers, e.g. Ma et al found 78 out of 142 driver genes in pediatric cancers undetected in adult pan-cancer studies (Ma 2018). Gröbner et al found single putative driver genes to be more specific of cancer type in pediatric cancers, as compared to adult cancers (Gröbner 2018). Ma et al also found more than half of detected driver genes to be specific for cancer types in a study across six different (mainly blood) cancers, e.g. TAL1 in T cell acute lymphoblastic leukemia (T-ALL) and ALK for neuroblastoma (NBL) (Ma 2018). Also, each subtype in for example MBB-SHH has specific pathways affected (Skowron 2021).

Gröbner et al found the RTK/MAPK signaling pathway, transcriptional regulation, cell cycle control and DNA repair as the main pathways affected by alterations (Gröbner 2018). Putative driver genes related to epigenetic regulation are on average more often mutated in some pediatric cancers than in adult cancers (Huether 2014, Gröbner 2018). Some alterations are very specific for pediatric cancer types, like the loss-of-function mutations of genes encoding histone H3 isoforms in diffuse intrinsic pontine gliomas (DIPG) (Wu 2012).

Non-coding and regulation

Non-coding mutations in regulatory parts of the genome upstream of coding genes can also alter gene expression. Some of these consist of enhancer- or promotor-hijacking events, which may also be caused by structural variants. For example, both Valentijn et al and Gröbner et al found recurrent alterations in TERT promoter regions in neuroblastomas (Valentijn 2015; Gröbner 2018). They seem to play a larger role than in adult cancers, for instance Reyna rarely found recurrent non-coding mutations in a large study of more than 2500 whole cancer genomes from 27 adult tumor types (Reyna 2020).

Structural variants and copy number variations

Next to driver mutations caused by SNVs and indels, structural variants (SV) and copy number variations (CNVs) can also result in cancer promoting events. Gröbner et al found recurrent copy-number changes in regions containing known oncogenes (Gröbner 2018). CNVs and SVs were found to be 62% of the events in a study of (mainly blood) pediatric cancers (Ma 2018). Ciriello et al created a crude classification based on these various functional events: M-class (SNV-dominated) or C-class (SV/CNV-dominated) (Ciriello 2013). Gröbner et al found that 55% of pediatric tumors were solely categorized in the M-class (small mutations) or the SC-class (structural/CNV), 27% were dominated by one type but also contained the other type, 8% were completely ambiguous and for 10% events could not be established (Gröbner 2018).

Gene fusions

Translocations resulting from SVs can result in gene fusions like the fused BCR-ABL oncogene in chronic myelogenous leukemia (CML) or incidentally can separate tumor suppressor genes from their promoters, both resulting in changes in gene expression and pathogenic effects (Vogelstein 2013). McLeod et al discussed targetable gene fusions in pediatric cancer with kinase-activating examples, e.g. in Philadelphia chromosome–like ALL (Ph-ALL) (McLeod

2021; Roberts 2014). Fusions arise from structural variants. Van Belzen et al found 165 high-confidence tumor-specific gene fusions in a pediatric data set of 130 patients with a method combining information from WGS and RNA-seq data. They argued that it is hard to predict the functional effects of these fusions and relate those to gene expression changes, since they often occur in genomic regions with large disruptions and copy number alterations which may also affect gene expression (Van Belzen 2021b). The medium number SVs varies wildly across pediatric cancer types with most SVs in cancer types having high genomic instability, like high grade glioma (HGG), Adenoid cystic carcinoma (ACC) and osteosarcoma (OS) among others (Gröbner 2018).

Summary differences and consequences

In summary, adult and pediatric cancers share only about 45% of driver genes and pediatric cancer consists of specific cancer types. Therefore findings for genetic interactions for adult cancers cannot be used for pediatric cancers. Additionally, there is a lower mutational load in pediatric cancers, while specific mutations and epigenetic changes during developmental and structural variants resulting in clinically relevant gene fusions appear to play a larger role in pediatric cancers than in adult cancers (Ma 2018, Gröbner 2018). Consequently, to study genetic interactions in pediatric cancer it is important to include not only somatic SNVs and indels, but also other relevant, and likely functional, events such as germline mutations, gene fusions, non-coding mutations and methylation data.

Results in pediatric cancer

We know turn to the results on genetic interaction in pediatric cancers based on the aforementioned pediatric data sets. Ma et al found 304 known and new (both CO and ME) genetic interactions in the TARGET data set (Ma 2018). However, they may have overestimated the number of CO interactions due to using Fisher's exact test. Gröbner et al found that most genes were mutually exclusive across cancer types in a pan-cancer analysis of just the significantly mutated genes in the DKFZ data set. They argued that they demonstrated specificity of single driver genes in pediatric cancers, and this differed from more co-occurring alterations in adult cancers. They used matrix permutation methods that kept the margins per gene and per sample constant (Gröbner 2018).

Daub et al applied the permutations method by Park et al (Park 2015), WeSME by Kim et al (Kim 2017) and DISCOVER by Canisius et al (Canisius 2016) to the two aforementioned pediatric cancer sets DFKZ and TARGET containing 23 cancer types. Unlike the existing analyses, they included all genes in their analysis (not only significantly mutated genes) and compared the WeSME results to permutation methods. They found 15 co-occurring and 27 mutually exclusive candidate gene pairs and 29% of all gene pairs specific for cancer subtype (and 44% for co-occurring gene pairs) (Daub 2020). 77% percent of the candidate pairs were confirmed with the DISCOVER method.

Daub et al further inspected possible biological reasons for the co-occurrence patterns, often missing from computational screen studies. They concluded that tumor type, pathway epistasis and to a lesser extent synthetic lethality are the most likely reasons for these patterns. An example of intra-pathway epistasis is the ME-pair PTCH1 and SMO, both part of

the Hedgehog signaling pathway. The ME pair NRAS-WT1 in AML was found to be a SL candidate.

There are also results from cancer type specific studies and we name two examples. Valentijn et al found in a cancer type specific study of neuroblastomas that (non-coding) TERT alterations, ATRX deletions and MYCN amplifications were mainly mutually exclusive between themselves, characterizing various subtypes of neuroblastoma (Valentijn 2015). Another example is the study by Skowron et al on Sonic Hedgehog Medulloblastoma (MBL-SHH): they included SNVs, CNVs, fusions and methylation patterns and applied DISCOVER in their search for interactions. Interestingly, they also found that diverse molecular mechanisms converged on recurrently mutated genes or pathways. For example, both transcriptional silencing by methylation as somatic mutations could lead to loss of expression of PAX6 (Skowron 2021). This highlights the importance of including different types of alterations.

The search for gene pairs does not seem to approach saturation yet. This may be shown by down sampling and statistical analyses. Daub et al showed by down sampling that reduced sample sizes are a possible cause for not being able to reproduce findings in smaller data sets (Daub 2020). Due to the small statistical power, we expect that many relevant interactions have not yet been found.

Discussion

Choice of data structure (aggregation level, cancer type, mutational signature etc.)

We have already discussed the effect of the granularity of alterations (genes, modules, pathways) on results and statistical power. In this part we reflect on the choice of structure of samples. Daub et al found that genetic interactions were mostly specific for tumor type or tissue type (Daub 2020). Let us look at two examples. First, suppose we try to find co-occurring (CO) mutations that are characteristic of a certain tumor type. Second, suppose we try to find mutually exclusive (ME) mutations in genes or pathways that reflect two kinds of subtypes within the tumor type, with underlying redundant pathway mechanisms (but both pathways are still very specific for this tumor type). Put otherwise, we aim to find ME drivers with redundancy as a cause that can define different putative cancer subtypes within the same tissue type (Kim 2017). In both cases, when we do a per tumor-type test we will have a better signal in the data than if we add samples from other tumor types in a pan-cancer test that would add noise to the data and dilute the signal. Additionally, if we would have divided the data further based on the *sub*types, than we would not have been able to find ME pairs in the second example, because the data of the other subtype are not included. From these examples we can see that the detection of pairs depends on a combination of the data structure (the way we subset the data), the number of samples, the type of interaction (CO, ME) and the underlying cause (subtypes, redundancy, etc.). Pan-cancer tests use more of the available samples but do not per definition result in a higher statistical power.

In general, gene interactions may be correlated to all kind of biological contexts such as tissue, cancer type or even cancer subtype, developmental stage or any *internal* cellular gene expression state. To detect such correlations in general, researchers do a confounding analysis. They try to find enrichment of gene pairs in subsets of the samples corresponding to the chosen data structure. Canisius et al argued that the findings of interactions depend

largely on this chosen data structure or what they named sources of heterogeneity (Canisius 2016). As we have seen in the examples, this has different effects on detection depending on the kind of interaction and underlying biological mechanism. Related to this, Kim et al introduced three different classes of mutual exclusivity (ME): within tissue type, across tissue type (common) and between tissue type exclusivity (Kim 2015). Across tissue type ME gene pairs that are not detected as within tissue type ME, may be detected as such with larger data sets, depending on the size of the data sets.

This results in circularity of the results depending on our hypotheses: by choosing the right underlying data structure for the kind of interaction we are looking for, we may find more gene pairs. To complicate things further, the direction of the interaction may also be dependent on the cancer subtype. Some pairs may be cooperative in one tumor, but result in synthetic sickness in another because of a different genetic background. In general, ME drivers with synthetic lethality as a cause may correlate with a certain context, as in induced synthetic lethality, where the SL depends on the internal state of the cell or other conditions outside the cell, such as medium or drugs. As a final example, we could address the biological question how genetic interactions interact with processes that result in mutational signatures. Mutational signatures potentially reflect mutagenic processes in the cell and may reveal causes of cancer (Alexandrov 2013). By choosing a data structure with samples assigned to one or more mutational signatures, the interaction of gene pairs and possible causal relations may be revealed.

We also have to be aware that intra-tumor heterogeneity can be a confounder. Resolving clones in the same biological sample through single-cell genomics/transcriptomics may counteract this and even help to improve signal when it enables discrimination between different clones. However, this is a very expensive endeavor.

To conclude, every choice of data structure implies making assumptions, and these should be made explicit and motivated as being relevant for the biological question.

Models and validation

Computational and experimental screens can point to a possible genetic interaction between genes. A difficulty shared with all methods in molecular biology that apply backward modeling to data mining in large omics data sets is to make a connection between the statistical findings and possible mechanisms (Gunawardena 2014). Various mutational loads, technical artifacts or chosen data structure may result in false positives. We should also be aware that (true positive) interactions may depend on the biological context, that may for example show up as different statistical results of co-occurrence tests when choosing various data structures. To understand the interaction and translate it into clinical applications we would prefer a forward model: we like to know about the exact molecular mechanisms of the interaction and predict its dependence on conditions like cell state. At the least, gene pairs need to be functionally validated in an experiment in model organisms/systems, for example in human and cancer cell lines, organoids or xenografts. Since mutation in both genes of a ME pair may have either a selective disadvantage or no selective *dis*advantage, validation is needed to know if synthetic lethality applies (O'Neil 2017). Synthetic lethal pairs can be validated by specific (knock out) gene editing with CRISPR-Cas9 methods. Other more subtle

interactions resulting in a change of gene expression are harder to validate. Even with validation in model systems, the genetic interactions still do not have to exist in patients, due to the dependence on the metabolic state or developmental stage of the tumor cells or on the cellular environment. A gene pair can have an interaction with the *external* environment of the cell (e.g. extracellular matrix (ECM)). When such a dependency can be revealed one speaks of induced/conditional synthetic lethality. This is a broad concern, and O'Neil et al describe some of the challenges, like tumor-associated conditions, the effect of therapeutic agents on the genetic interaction network and differential sensitivity between tumor and non-tumor cells (O'Neil 2017).

Data availability and prioritization

We focus on how computational screens can be improved. Since pediatric cancers have a smaller incidence and a lower mutational burden, the available data sets are generally smaller and samples have a lower mutational load and thus statistical tests using these data sets have less statistical power. Both the identification of single driver genes as the detection of gene pairs depend on identification of variants, which may be hampered by technical and biological reasons like mapping ambiguities, tumor purity or subclones in the tumor among others. Besides, intra-tumor heterogeneity has a more specific effect on identifying gene pairs depending on the type of interaction. A high level of tumor heterogeneity in one cancer type results in generation of false-positive positive interactions and false-negative mutually exclusive (negative) interactions because gene mutations in different individual cells would be incorrectly detected as co-occurring (the reads of different cells of different clones would be sequenced together) (Park 2015). Furthermore, SL interactions are rare and need a lot of samples to be detected.

What data can we easily include in our analysis on gene level? SVs and CNVs seem to promise the largest addition to count matrices, especially if we focus on gene expression loss, but come with some challenges. First, the identification of tumor-specific structural variants is far from trivial and may also origin from mosaic variation in healthy tissue (Van Belzen 2021a). Second, Vogelstein argued that it is difficult to identify driver alterations when large parts of DNA containing multiple genes are either duplicated or deleted (Vogelstein 2013). Indeed, including counts for all lost genes in such a long part may dilute the signal for detection of genetic interactions.

Also, non-coding mutations and germline mutations seem to have a larger role in pediatric cancers than in adult cancers and can easily be included.
Gene fusions also seem to make a large contribution to cancer progression. However, functional prediction and the causal relation with gene expression can be difficult to detect, especially if copy number alterations are found in the same genomic regions where the breakpoints resulting in gene fusions take place. Therefore, we should also try to avoid double counting of affected genes in a sample. Another issue is that gene fusions - per definition - have a different granular level than that of single genes. One should decide how to count the contribution of the genes participating in the gene fusions in count matrices. This is easier for reoccurring genes than for incidental partners in fusions.

Finally, methylation data can be a promising addition, since the epigenetic phenomena have a large contribution in pediatric cancer. Vogelstein et al noticed that "epi-driver" genes (altered by epigenetic mechanisms) are harder to identify than "mut-driver" genes (by DNA aberrations) (Vogelstein 2013). Methylation can be transient and samples can be heterogeneous in methylation patterns (Wenger 2019). However, methylation profiling is used ever more for characterization of cancers in clinical and research settings, mainly by use of either bisulfite (BS) sequencing methods or multiplexed probe-base detection methods. At least observed patterns should be confirmed by RNA-seq data.

Data integration

Due to the comparable low abundance of data on pediatric cancers, an open question is how to get the most statistical power out of the available data and how to integrate data of different types, e.g. how to combine epigenetic data sets and data sets of various mutational types, like SNVs, structural variants and gene fusions. We discuss three ways to integrate the data, i) enrichment/horizontal integration: taking all features of multi-omics data together in a single feature space or ii) vertical integration: combining various data sets of the same kind and iii) model integration: combining the predictive power of various models in heuristic, statistical or graph-based methods.

Let us first consider the single feature space and see how we would define a feature. In most methods any likely functional event on gene level is taken into account, accepting the possibility that some functional events may be complete gene loss or may have any of the other quantitative effects discussed at the beginning of this review. Detail is lost by not specifying the direction and quantity of the alterations and this may hamper the detection of gene pairs and connection to underlying mechanisms. To counteract the possibility of averaging out these different effects, one could alternatively only combine all alterations resulting in complete gene (expression) loss, whether it would result from SNVs, structural variants or methylation. A possible challenge can be that data sets may not be completely overlapping and one needs a strategy for missing data, either by only using the rows of the intersection of the data set or fill in the data by less precise variant prediction methods or machine learning methods.

We may also try to increase statistical power by vertical integration: combining various data sets of the same kind, for example those only containing SNVs or only tumor-specific structural variants. Ideally one would like to apply identical pipelines to multiple data sets. However, methods may depend on different data. An example is the use of RNA expression data to confirm copy number alterations and gene fusions in tumor-specific SV calling (van Belzen 2021b). Other pipelines for calling structural variants may not include RNA-seq data and thus methods cannot be applied to all data sets. It can often be computationally expensive to reprocess the raw data. When raw data are not available or cannot be reprocessed due to various dependencies, the resulting binary gene/sample matrices have different underlying filters and kinds of data and cannot easily be combined. Daub et al suggested to permutate these matrices per cancer type within data sets and thus determine false discovery rates (Daub 2020). In this way we combine the significance of the predictions based on different data sets and this is more like model integration, which we further discuss

below. This emphasizes the importance of open data and open methods: collaborations between institutes with data sharing platforms with standardization of pipelines such as in the Pan-Cancer analysis of whole genomes (PCAWG) consortium for adult cancers (Reyna 2020) or the CBTTC data available in the Gabriella Miller Kids First Data Resource Center (KF-DRC, https://kidsfirstdrc.org).

Finally, model integration can be applied by using various methods on i) the same data ii) multiple data sets of the same types of data or iii) of different types of data. Using multiple models on the same data may help in confidence, may avoid false positives, but should not result in large additions and therefore have limited added value in finding *new* interactions. Model integration can be done heuristically by using majority voting of multiple models. To be more precise, we can also quantitatively combine the confidence or significance scores (e.g. p-values) of the predictions of multiple models. As we have seen in the experimental methods, this can also be applied to findings in cell lines and model organisms/systems. However, it is a challenge to decide how to weight these: i) genetic interactions may depend on *internal* (RNA expression, metabolic state) and *external* conditions (e.g. extra cellular matrix (ECM)) and ii) orthologues in other species with a complete different genetic background and regulatory network may have a different effect. In some methods results of phylogenetic research are also used as a filter for candidate gene pairs. However, this does not help with an extensive genome-wide search.

Conclusion

Interactions between genes in pediatric cancer can be identified using experimental and computational methods. The statistical power of tests to detect these gene pairs has been a recurrent theme in this review. Many aspects hamper power. There are many alterations in genes and pathways that result in similar functional effects and pediatric cancer types or their subtypes are often rare. This suggests a long tail of rare interactions that can only be revealed with enough statistical power. Detection of those rare interactions can be improved by adding and integrating data, choosing the right data structure and improving on methods. In adult cancer we have seen that applying the knowledge of synthetic lethality to the clinic takes a lot of time, and as a first step results of screens need to be carefully validated and mechanisms understood. Although the genome-wide search for genetic interactions in pediatric cancer types is only in its initial phase, incorporation of new data and new methods is promising and can result in new therapies.

References

(1) Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013 August 22;500(7463):415-421.

(2) Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. Cell 2011 April 01;145(1):30-38.

(3) Boettcher M, Tian R, Blau JA, Markegard E, Wagner RT, Wu D, et al. Dual gene activation and knockout screen reveals directional dependencies in genetic networks. Nat Biotechnol 2018 February 01;36(2):170-178.

(4) Bolouri H, Farrar JE, Triche T, Ries RE, Lim EL, Alonzo TA, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nat Med 2018 January 01;24(1):103-112.

(5) Brody LC. Treating cancer by targeting a weakness. N Engl J Med 2005 September 01;353(9):949-950.

(6) Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature 2005 April 14;434(7035):913-917.

(7) Byrjalsen A, Diets IJ, Bakhuizen J, Hansen TVO, Schmiegelow K, Gerdes AM, et al. Selection criteria for assembling a pediatric cancer predisposition syndrome gene panel. Fam Cancer 2021 June 01.

(8) Calandrini C, Schutgens F, Oka R, Margaritis T, Candelli T, Mathijsen L, et al. An organoid biobank for childhood kidney cancers that captures disease and tissue heterogeneity. Nat Commun 2020 March 11;11(1):1310-6.

(9) Canisius S, Martens JW, Wessels LF. A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol 2016 December 16;17(1):261-x.

(10) Capper D, Stichel D, Sahm F, Jones DTW, Schrimpf D, Sill M, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. Acta Neuropathol 2018 August 01;136(2):181-210.

(11) Carter H, Marty R, Hofree M, Gross AM, Jensen J, Fisch KM, et al. Interaction Landscape of Inherited Polymorphisms with Somatic Events in Cancer. Cancer Discov 2017 April 01;7(4):410-423.

(12) Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. Nat Genet 2013 October 01;45(10):1127-1133.

(13) Cui X, Han L, Liu Y, Li Y, Sun W, Song B, et al. siGCD: a web server to explore survival interaction of genes, cells and drugs in human cancers. Brief Bioinform 2021 April 05.

(14) Custers L, Khabirova E, Coorens THH, Oliver TRW, Calandrini C, Young MD, et al. Somatic mutations and single-cell transcriptomes reveal the root of malignant rhabdoid tumours. Nat Commun 2021 March 03;12(1):1407-6.

(15) Daub JT, Amini S, Kersjes DJE, Ma X, Jäger N, Zhang J, et al. A comprehensive map of genetic interactions in childhood cancer reveals multiple underlying biological mechanisms. bioRxiv 2020:2020.11.17.385120.

(16) Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al. Mutations of the BRAF gene in human cancer. Nature 2002 June 27;417(6892):949-954.

(17) Deng Y, Luo S, Deng C, Luo T, Yin W, Zhang H, et al. Identifying mutual exclusivity across cancer genomes: computational approaches to discover genetic interaction and reveal tumor vulnerability. Brief Bioinform 2019 January 18;20(1):254-266.

(18) Dhanjal JK, Radhakrishnan N, Sundar D. Identifying synthetic lethal targets using CRISPR/Cas9 system. Methods 2017 December 01;131:66-73.

(19) Domingo J, Baeza-Centurion P, Lehner B. The Causes and Consequences of Genetic Interactions (Epistasis). Annu Rev Genomics Hum Genet 2019 August 31;20:433-460.

(20) Downing JR, Wilson RK, ZhanFAU - Mardis E, Mardis ER, Pui CH, Ding L, et al. The Pediatric Cancer Genome Project. Nat Genet 2012 May 29;44(6):619-622.

(21) Drost J, van Jaarsveld RH, Ponsioen B, Zimberlin C, van Boxtel R, Buijs A, et al. Sequential cancer mutations in cultured human intestinal stem cells. Nature 2015 May 07;521(7550):43-47.

(22) Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol 2011;6:479-507.

(23) Gunawardena J. Models in biology: 'accurate descriptions of our pathetic thinking'. BMC Biol 2014 April 30;12:29-29.

(24) Hanemaaijer ES, Margaritis T, Sanders K, Bos FL, Candelli T, Al-Saati H, et al. Single-cell atlas of developing murine adrenal gland reveals relation of Schwann cell precursor signature to neuroblastoma phenotype. Proc Natl Acad Sci U S A 2021 February 02;118(5):10.1073/pnas.2022350118.

(25) Huang A, Garraway LA, Ashworth A, Weber B. Synthetic lethality as an engine for cancer drug target discovery. Nat Rev Drug Discov 2020 January 01;19(1):23-38.

(26) Huang KL, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. Cell 2018 April 05;173(2):355-370.e14.

(27) Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun 2014 April 08;5:3630.

(28) Izzo F, Lee SC, Poran A, Chaligne R, Gaiti F, Gross B, et al. DNA methylation disruption reshapes the hematopoietic differentiation landscape. Nat Genet 2020 April 01;52(4):378-387.

(29) Kafaie S, Xu L, Hu T. Statistical methods with exhaustive search in the identification of gene-gene interactions for colorectal cancer. Genet Epidemiol 2021 March 01;45(2):222-234.

(30) Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013 October 17;502(7471):333-339.

(31) Kildisiute G, Kholosy WM, Young MD, Roberts K, Elmentaite R, van Hooff SR, et al. Tumor to normal single-cell mRNA comparisons reveal a pan-neuroblastoma cancer cell. Sci Adv 2021 February 05;7(6):10.1126/sciadv.abd3311. Print 2021 Feb.

(32) Kim JW, Botvinnik OB, Abudayyeh O, Birger C, Rosenbluh J, Shrestha Y, et al. Characterizing genomic alterations in cancer by complementary functional associations. Nat Biotechnol 2016 May 01;34(5):539-546.

(33) Kim YA, Cho DY, Dao P, Przytycka TM. MEMCover: integrated analysis of mutual exclusivity and functional network reveals dysregulated pathways across multiple cancer types. Bioinformatics 2015 June 15;31(12):284.

(34) Kim YA, Madan S, Przytycka TM. WeSME: uncovering mutual exclusivity of cancer drivers and beyond. Bioinformatics 2017 March 15;33(6):814-821.

(35) Laufer C, Fischer B, Billmann M, Huber W, Boutros M. Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. Nat Methods 2013 May 01;10(5):427-431.

(36) Lee JS, Nair NU, Dinstag G, Chapman L, Chung Y, Wang K, et al. Synthetic lethality-mediated precision oncology via the tumor transcriptome. Cell 2021 April 29;184(9):2487-2502.e13.

(37) Lehner B. Molecular mechanisms of epistasis within and between genes. Trends Genet 2011 August 01;27(8):323-331.

(38) Leiserson MD, Reyna MA, Raphael BJ. A weighted exact test for mutually exclusive mutations in cancer. Bioinformatics 2016 September 01;32(17):i736-i745.

(39) Leiserson MD, Vandin F, Wu HT, Dobson JR, Eldridge JV, Thomas JL, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat Genet 2015 February 01;47(2):106-114.

(40) Locke WJ, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KYC, et al. DNA Methylation Cancer Biomarkers: Translation to the Clinic. Front Genet 2019 November 14;10:1150.

(41) Magen A, Das Sahu A, Lee JS, Sharmin M, Lugo A, Gutkind JS, et al. Beyond Synthetic Lethality: Charting the Landscape of Pairwise Gene Expression States Associated with Survival in Cancer. Cell Rep 2019 July 23;28(4):938-948.e6.

(42) Mani R, St Onge RP, Hartman JL, Giaever G, Roth FP. Defining genetic interaction. Proc Natl Acad Sci U S A 2008 March 04;105(9):3461-3466.

(43) McLeod C, Gout AM, Zhou X, Thrasher A, Rahbarinia D, Brady SW, et al. St. Jude Cloud: A Pediatric Cancer Genomic Data-Sharing Ecosystem. Cancer Discov 2021 May 01;11(5):1082-1099.

(44) McLornan DP, List A. Applying synthetic lethality for the selective targeting of cancer. N Engl J Med 2014 October 30;371(18):1725-1735.

(45) Najm FJ, Strand C, Donovan KF, Hegde M, Sanson KR, Vaimberg EW, et al. Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. Nat Biotechnol 2018 February 01;36(2):179-189.

(46) Niemeyer CM, Flotho C. Juvenile myelomonocytic leukemia: who's the driver at the wheel? Blood 2019;133(10):1060-1070.

(47) O'Neil NJ, Bailey ML, Hieter P. Synthetic lethality and cancer. Nat Rev Genet 2017 October 01;18(10):613-623.

(48) Park S, Lehner B. Cancer type-dependent genetic interactions between cancer driver alterations indicate plasticity of epistasis across cell types. Mol Syst Biol 2015 July 30;11(7):824.

(49) Parsons DW, Roy A, Yang Y, Wang T, Scollon S, Bergstrom K, et al. Diagnostic Yield of Clinical Tumor and Germline Whole-Exome Sequencing for Children With Solid Tumors. JAMA Oncol 2016 May 01;2(5):616-624.

(50) Reyna MA, Haan D, Paczkowska M, Verbeke LPC, Vazquez M, Kahraman A, et al. Pathway and network analysis of more than 2500 whole cancer genomes. Nat Commun 2020 February 05;11(1):729-0.

(51) Roberts KG, Li Y, Payne-Turner D, Harvey RC, Yang YL, Pei D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. N Engl J Med 2014 September 11;371(11):1005-1015.

(52) Skowron P, Farooq H, Cavalli FMG, Morrissy AS, Ly M, Hendrikse LD, et al. The transcriptional landscape of Shh medulloblastoma. Nat Commun 2021 March 19;12(1):1749-0.

(53) Solman M, Blokzijl-Franke S, Yan C, Yang Q, Kamel SM, Bakkers J, et al. Inflammation potentiates JMML-like blood defects in Shp2 mutant Noonan syndrome. bioRxiv 2020:2020.09.10.289090.

(54) Sweet-Cordero EA, Biegel JA. The genomic landscape of pediatric cancers: Implications for diagnosis and treatment. Science 2019;363(6432):1170-1175.

(55) Topatana W, Juengpanich S, Li S, Cao J, Hu J, Lee J, et al. Advances in synthetic lethality for cancer therapy: cellular mechanism and clinical translation. J Hematol Oncol 2020 September 03;13(1):118-5.

(56) Underhill C, Toulmonde M, Bonnefoi H. A review of PARP inhibitors: from bench to bedside. Ann Oncol 2011 February 01;22(2):268-279.

(57) Unni AM, Lockwood WW, Zejnullahu K, Lee-Lin SQ, Varmus H. Evidence that synthetic lethality underlies the mutual exclusivity of oncogenic KRAS and EGFR mutations in lung adenocarcinoma. Elife 2015 June 05;4:e06907.

(58) Valentijn LJ, Koster J, Zwijnenburg DA, Hasselt NE, van Sluis P, Volckmann R, et al. TERT rearrangements are frequent in neuroblastoma and identify aggressive tumors. Nat Genet 2015 December 01;47(12):1411-1414.

(59) van Belzen, Ianthe A. E. M., Cai C, Tuil Mv, Badloe S, Strengman E, Janse A, et al. Systematic discovery of gene fusions in pediatric cancer by integrating RNA-seq and WGS. bioRxiv 2021:2021.08.31.458342.

(60) van Belzen, Ianthe A. E. M., Schönhuth A, Kemmeren P, Hehir-Kwa J. Structural variant detection in cancer genomes: computational challenges and perspectives for precision oncology. npj Precision Oncology 2021;5(1):15.

(61) Vandin F, Upfal E, Raphael BJ. De novo discovery of mutated driver pathways in cancer. Genome Res 2012 February 01;22(2):375-385.

(62) Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science 2013 March 29;339(6127):1546-1558.

(63) Wang Y, Xiao M, Chen X, Chen L, Xu Y, Lv L, et al. WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. Mol Cell 2015 February 19;57(4):662-673.

(64) Weidema ME, van de Geer E, Koelsche C, Desar IME, Kemmeren P, Hillebrandt-Roeffen MHS, et al. DNA Methylation Profiling Identifies Distinct Clusters in Angiosarcomas. Clin Cancer Res 2020 January 01;26(1):93-100.

(65) Wenger A, Ferreyra Vega S, Kling T, Bontell TO, Jakola AS, Caren H. Intratumor DNA methylation heterogeneity in glioblastoma: implications for DNA methylation-based classification. Neuro Oncol 2019 May 06;21(5):616-627.

(66) Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfort J, et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat Genet 2012 January 29;44(3):251-253.

(67) Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. N Engl J Med 2015 December 10;373(24):2336-2346.