GORDION BIOSCIENCE

POSITIONING OF WRN AS AN ONCOLOGY DRUG TARGET

SCORE :

Successful target-like properties 2/25

The proprietary AI model was trained to differentiate successful drug targets from those that failed clinical trials
Publicly available datasets massively underestimated the frequency of deleterious WRN mutations in patient tumor samples, creating an inflated perception of WRN attractiveness as a therapeutic target.

Translatability of pre-clinical models 3/25

The technology evaluates whether dependencies observed in cell lines are present in real patient tumor data
While WRN dependency in MSI-H is detectable in patient samples, cell line models substantially overestimated its importance. Other dependencies observed in CRISPR screens were detected but are weak in patient data.

Clinical response rates in the primary indication 2/25

The AI model measures how critical for the tumor is the activity of drug target
WRN appears important but not critical for MSI-H tumors. Based on the strength of the signal in patient data, as a monotherapy WRNi are unlikely to succeed.

Market expansion and competitive advantage over other drugs 1/25

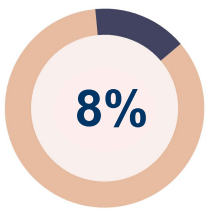
The AI model compares WRN with other drug targets in possible indications
Multiple other targets demonstrated stronger drug target-like properties and higher potential clinical utility than WRN helicase in tested indications.

Conclusions & recommendations

Patient data, in contrast to preclinical models, suggest that inhibition of WRN will not affect tumor growth.

WRN inhibitors are unlikely to succeed in clinical trials.

PROBABILITY OF SUCCESS:



ANALYSIS OF WRN POTENTIAL AS A DRUG TARGET

Summary

WRN (Werner Syndrome helicase) is widely explored as a drug target based on several CRISPR functional genomics screens on cell line panels.

Using Gordion technology, the WRN potential as a drug target was assessed. Several key observations informing the decision whether to continue the WRN program were made:

1. Publicly available databases and tools massively underestimated the frequency of WRN mutations in patient samples, inflating the attractiveness of WRN as a drug target.
2. Cell line models exaggerated the relevance of WRN as compared to real tumor data.
3. Only two tumor types (colorectal and ovarian) show the importance of WRN, yet the signal is an order of magnitude weaker than that observed for successful targets like PARP.
4. Consistent with basic research and CRISPR screens, several biomarkers were identified; however, the relationship between these biomarkers and WRN was relatively weak ($p > 0.05$).
5. The Platform identified an opportunity that could lead to reasonable clinical responses. The validation experiments were designed to test whether isochromosome 8 formation could lead to a larger therapeutic window for WRNi.

Conclusion & recommendation

Patient data suggest that inhibition of WRN **will not affect tumor growth.**

WRN inhibitors are unlikely to succeed in clinical trials.

Introduction to the technology

The concept of gene essentiality measurement in real-patient data

The Gordion Platform was designed to identify genes essential for cancer survival. The products of such genes constitute the best drug targets, which inhibition results in high clinical response rates.

Gene losses can lead to faster growth of a tumor cell, and such losses accumulate during tumor evolution (the best examples are tumor suppressor genes like p53 or PTEN). On the other hand, the loss of a gene essential for cancer survival leads to reduced fitness; therefore, such gene losses are eliminated during tumor evolution. Genes essential for tumor survival are frequently activated or amplified. Gordion's technology detects patterns of loss, activation, and amplification of any given gene and translates them into the probability of success for any drug target.

On the DNA level, the function of a gene can be lost in multiple ways, for example:

- the presence of a protein-altering single-nucleotide variant (SNV) or a short indel,
- by structural variants (SVs), affecting the gene sequence, promoters, or enhancers,
- by eliminating the entire copy of the gene by deletion.

Very few genes are lost completely (loss of all functional copies, bi-allelic loss) in sufficient numbers in tumors to facilitate robust statistical analysis. However, Gordion discovered that analysis of losses of a single gene copy (mono-allelic loss) directly in patient data, combined with the analysis of gene activations and amplifications, reveals the relevance of a gene for tumor survival.

Two different databases will be used for the analysis.

- 2,607 Whole-Genome Sequencing Datasets - a collection of sequences of whole cancer genomes and matched normal genomes from the same patient from breast, liver, lung, ovarian, and pancreatic cancers, with fully characterized genetic mutation profiles including single-nucleotide variants, in-dels, copy number alterations, and structural changes. The majority of the analyzed samples originated from refractory, metastatic tumors.
- 9,966 Copy Number Alterations Datasets combined with Whole-Exome Sequencing (CNA+WES). Samples originated from primary tumors.

Detailed information regarding databases is presented in the supplementary.

WRN as a drug target

WRN RecQ Like Helicase

HGNC: 12791
NCBI Gene: 7486
Ensembl: ENSG00000165392
OMIM®: 604611
UniProtKB/Swiss-Prot: Q14191

WRN, located on chromosome 8, encodes a RecQ helicase critical for DNA repair and genome stability (1,2). It is a possible synthetic lethal target in cancers with defective DNA damage repair (DDR), such as colorectal, endometrial, and ovarian tumors with DDR deficiencies (3). Inhibition of WRN in DDR-defective cancer cells induces DNA damage, genome instability, and selective tumor cell death independent of p53 status (3,4). In cell models, WRN downregulation or inhibition suppresses cell proliferation and promotes cell death, supporting its potential as a therapeutic target (4).

WRN is not accumulating point mutations

The cancer genome is burdened with a high number of mutations. However, their distribution is not random (5). Genes under positive selection pressure (oncogenes and tumor suppressors) tend to accumulate more point mutations than others, while those under negative selective pressure (essential genes) accumulate fewer (5). To analyze the accumulation of protein-altering mutations at chromosome 8, we used the Gene Mutational Load (GML) analysis.

The GML for a gene is calculated by counting protein-altering mutations (SNVs and indels) in each genome. The number of mutations is normalized by the mutational burden for the individual tumor (total mutations in that tumor's genome) to compensate for differences in mutation rates between the tumors, and by the length of the gene's coding sequence (6).

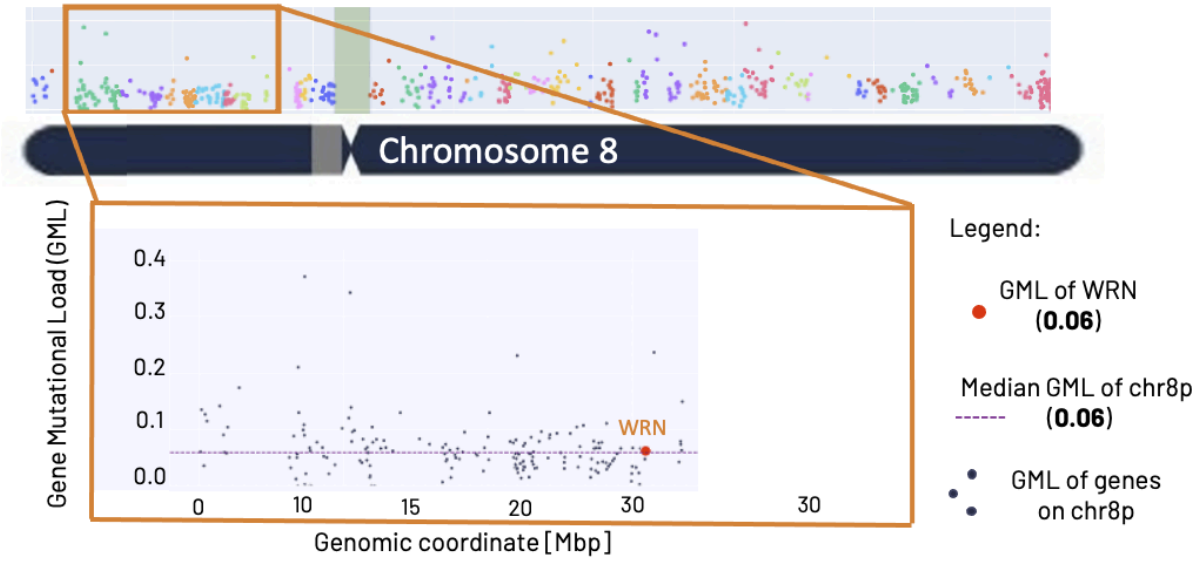


Fig. 1. *Analysis of Gene Mutational Load (GML), showing the burden of point mutations in genes on the p arm of chromosome 8. Each dot represents a distinct protein-coding gene, and the horizontal line is the median for all genes.*

GML analysis (Fig. 1) showed that WRN accumulates protein-altering mutations at a median rate of 0.06, which matches the average mutation rate of genes located on the p arm of chromosome 8. Typically, genes essential for tumor survival tend to accumulate fewer mutations than the average gene due to selective pressure to maintain their function. In the case of WRN, this pattern is not observed, suggesting there is no strong pressure to preserve its full function in tumors. This result raises the question of whether WRN is an attractive drug target.

The WRN gene is often lost

Both copies (bi-allellic loss) of a gene can be inactivated by two point mutations, two deletions, or a combination of a deletion and a point mutation. According to public databases and published results, the WRN gene is rarely bi-allelically lost in tumor samples (7,8).

In contrast with these results, Gordion observed that WRN is frequently bi-allelically lost in real tumors (Fig.2). For example, in breast cancer, according to the Cancer Genome Atlas (TCGA) genome collection (4), WRN is bi-allelically lost in 0.1% of samples. However, analysis performed by our proprietary ultra-sensitive algorithm for mutation detection revealed that in breast cancer samples, complete loss of WRN is dramatically more prevalent (10-fold higher) and exceeds 1%. Further investigation identified the mechanism behind this discrepancy: publicly available data and tools underestimate mutations that do not arise from point mutations, but instead are the result of two deletions removing both alleles.

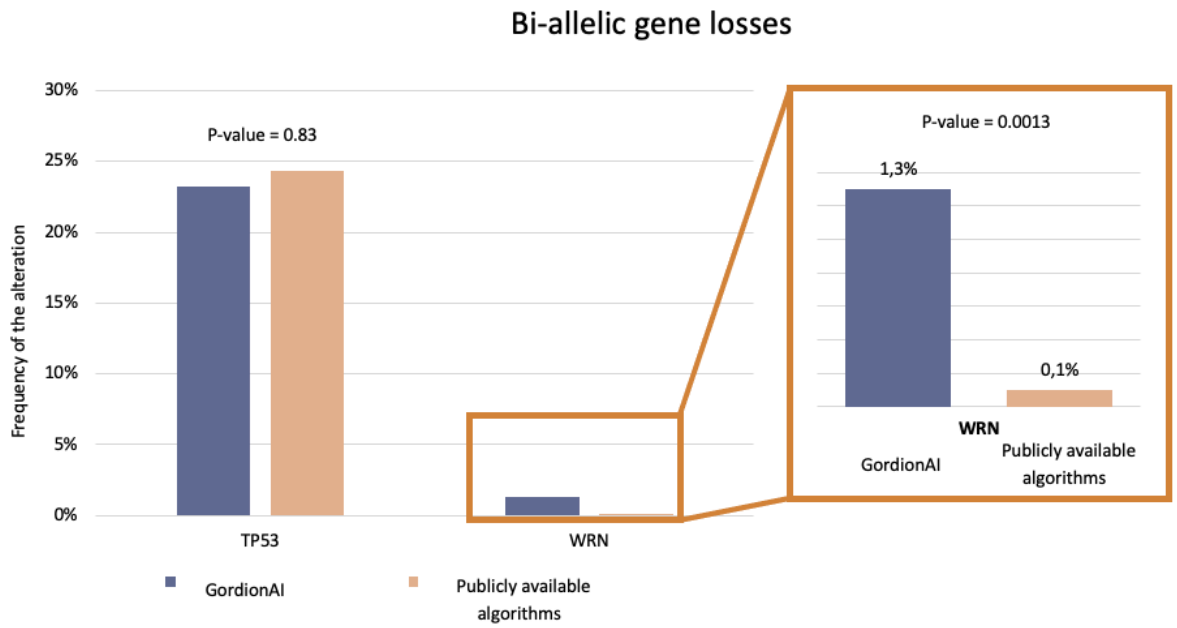


Fig. 2. *Comparative analysis of bi-allelic gene losses for WRN and TP53 in breast tumors, detected using publicly available algorithms used in genomic databases like TCGA and proprietary mutation detection AI technology developed by Gordion. Notably, the frequency of bi-allelic WRN loss is markedly elevated—over 10-fold higher than previously assessed—highlighting a dramatic enrichment compared to established reference datasets.*

As the majority of bi-allelic losses of WRN arise from two deletions spanning the entire WRN gene, these events are under-represented in standard analysis. As Gordion's AI algorithm was trained to detect this particular type of deletion, it allowed us to observe the true scale of WRN inactivation across tumor types.

Consistent with GML analysis from the previous section, this observation further questions the foundational assumption that WRN activity is critical for the survival of the real tumors.

WRN is rarely amplified

Genes essential for tumor survival are frequently amplified (5). A gene is considered amplified when it carries two or more additional copies above the mean tumor ploidy. WRN amplification was detected in 1.7% of tumors (47 out of 2,607 samples), higher than the median amplification rate across all genes (1.1%). However, WRN is located close to a strong driver of amplifications, and the data suggest that WRN is amplified as a passenger gene when the driver gene FGFR1 becomes amplified (9). These results suggest that tumors are unlikely to heavily rely on WRN, as it is a passenger of the amplification events.

Lack of Functional Compensation Among WRN Paralogs can lead to overestimation of a signal in the CRISPR screen

WRN has four paralogs: RECQL4, RECQL, BLM, and RECQL5, all members of the RecQ helicase family (10). We applied our AI algorithm to assess whether WRN deficiency correlates with upregulation of any of these paralogs. The analysis revealed no statistically significant association. This finding is consistent with the literature, which indicates that none of these paralogs can effectively compensate for WRN function, especially in the context of DNA repair and microsatellite instability (MSI) cancer cell survival (4,10,11). Each paralog has distinct substrate specificities, cellular localizations, and functional roles, preventing them from serving as functional backups for WRN loss.

The absence of functional compensation among WRN paralogs can lead to exaggeration of the WRN signal in CRISPR screens. The proprietary Gordion analysis comparing compensation mechanisms and relative signals in CRISPR versus patient data reveals that WRN is likely a false-positive signal in CRISPR screens.

Complete gene knockout in CRISPR screens eliminates gene function instantaneously, bypassing the cellular adaptation mechanisms and compensatory responses that tumors naturally develop over time to compensate for gene loss, creating artificially lethal dependencies that do not reflect the biological reality of tumor progression (12, 13). This explains why WRN performs exceptionally well in CRISPR screens and DepMap analyses as a synthetic lethal target—when genes lack compensatory paralogs, their functional disruption produces strong survival effects, making them appear as ideal therapeutic candidates. Indeed, CRISPR and RNAi screens consistently identify WRN as the top preferential dependency in MSI compared to microsatellite stable (MSS) cell lines, with 73% of MSI lines showing WRN dependency while other RecQ helicases show no such selective essentiality, further demonstrating the artificial nature of this dependency signal (1, 10). This conclusion is consistent with recent clinical results showing low clinical response rates.

Essentiality of WRN in different tumor types

Despite discouraging results from previous sections, Gordion performed an unbiased analysis of WRN potential as a drug target. In the concept of Synthetic Lethality, under specific conditions, an alteration affecting the Synthetic Lethal gene is not tolerated and leads to cell death. Following the same principle, we asked whether there are any conditions with a lower occurrence of alterations downregulating WRN. Gordion is using all alterations leading to gene downregulation (alterations including SNV, Indels, SV, and deletions, both mono-allelic and bi-allelic) to identify tissues, where the reduction of WRN activity via loss of one gene copy will lead to cell death. The rationale: if WRN accumulates fewer mutations downregulating WRN than other genes, it might suggest the increased essentiality of WRN.

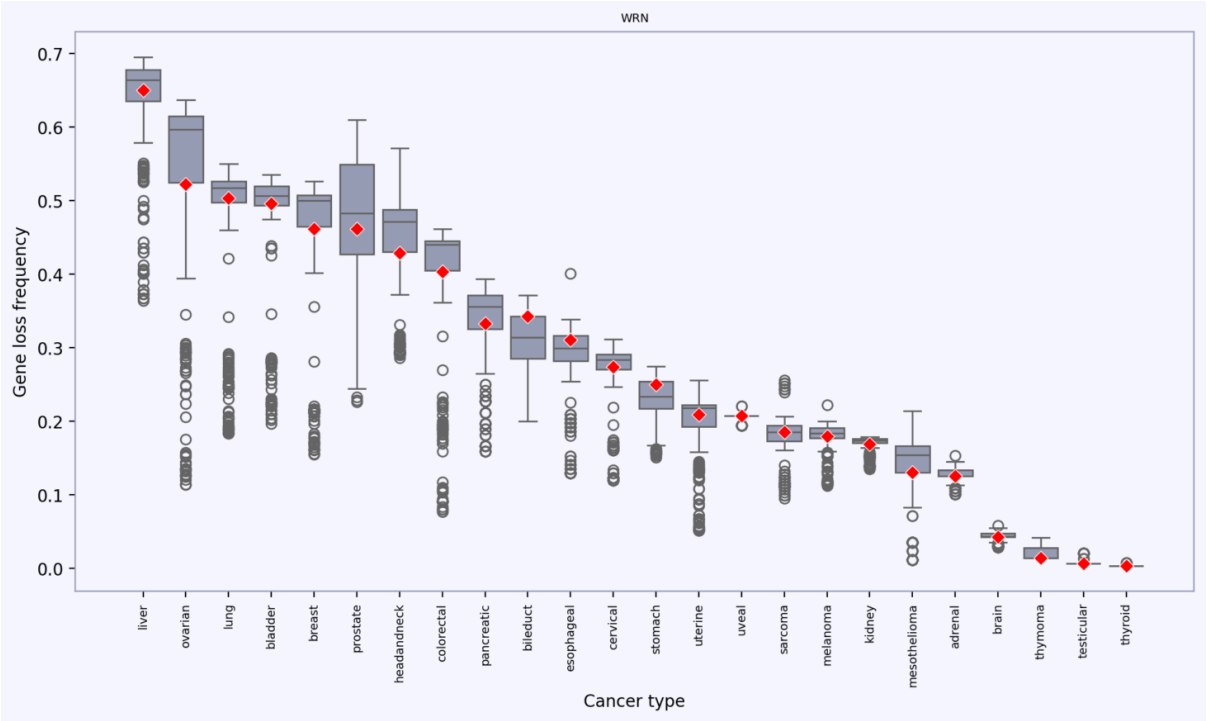


Figure 3. *Gene downregulation of WRN in different tumors is illustrated by the difference (delta) between the median gene loss across the chromosome arm (the horizontal line on the bar plots indicates) and WRN gene loss represented by a red diamond*

In several tumor types, such as ovarian, colorectal, breast, head and neck, prostate, and pancreatic cancers, the WRN gene is lost less frequently than the average gene, suggesting a potential selective advantage for retaining WRN in these contexts.

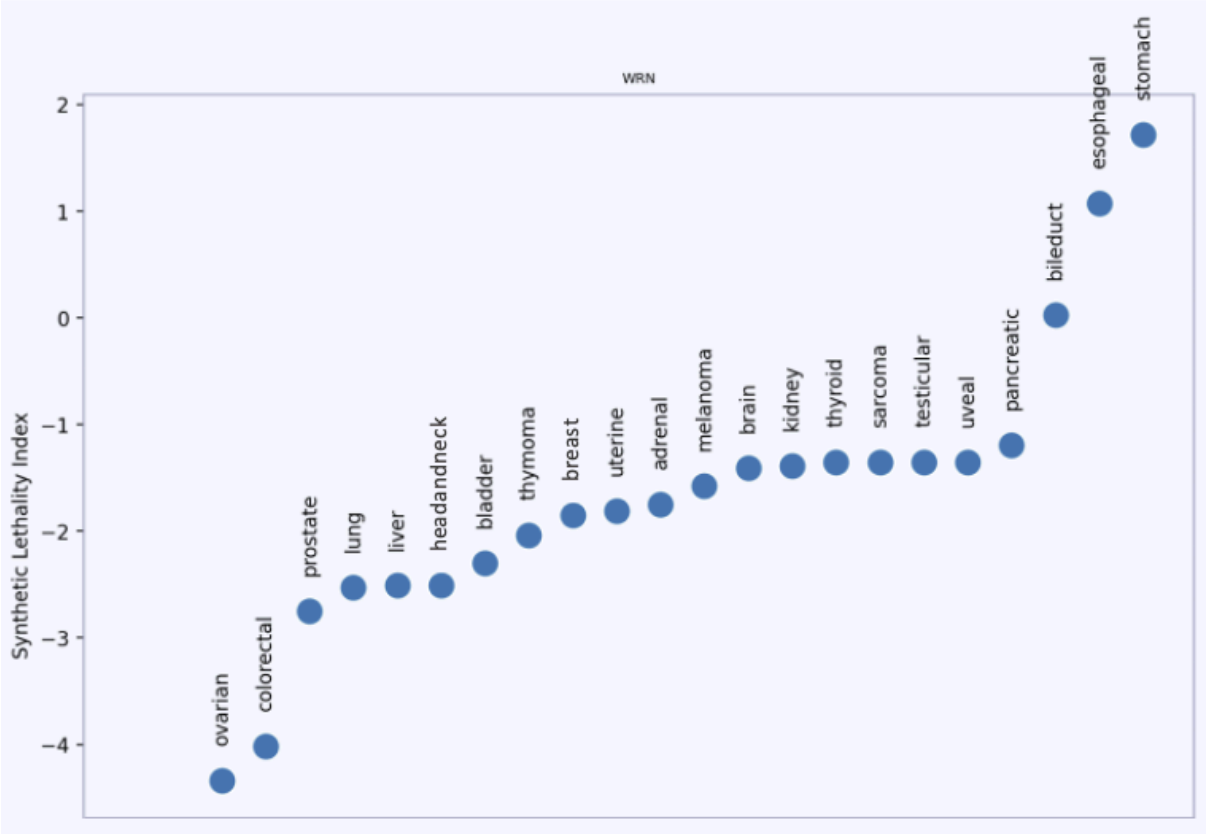


Figure 4. *The Synthetic Lethality Index integrates gene loss and amplification patterns to quantify WRN essentiality across different tumor types. Lower SL Index scores correspond to higher tissue dependency on WRN. This metric enables the identification of tumors most vulnerable to WRN-targeted inhibition.*

Essential genes are downregulated less frequently than average genes, but are amplified more often than the chromosome arm average. By integrating these two observables, Gordion developed a metric named the Synthetic Lethality Index (SL Index), which enables the identification of tissues where WRN essentiality is highest, with lower SL Index scores indicating greater essentiality. Notably, ovarian and colorectal tumors exhibit the highest WRN essentiality. In both ovarian and colorectal cancers, WRN loss occurs significantly less often than the average gene on the same chromosome arm, suggesting these tumor types are less tolerant of WRN loss. Both cancer types commonly harbor defects in DNA damage repair pathways, specifically deficiencies in homologous recombination (HR) or mismatch repair (MMR), which likely underlie their dependency on WRN function and the synthetic lethality relationship observed in CRISPR screens. This result suggests that there might be subcohorts of patients highly dependent on WRN where WRNi could deliver clinical efficacy.

Exploration of the genetic background for WRN inhibition in colorectal and ovarian cancers

We focused on colorectal and ovarian cancers because they exhibit the strongest dependency on WRN function, making them prime candidates for WRN-targeted therapies.

1. Increased essentiality of WRN in MSI-H colorectal patients

In colorectal cancer cohorts stratified by microsatellite instability (MSI) status, MSI-High (MSI-H) tumors showed a modest but consistent increase in dependency on WRN. Our analysis (Fig. 5) found that MSI-H colorectal tumors exhibit a reduction in the median frequency of WRN gene downregulation (from 41% to 20%), with no increased level of WRN gene amplification. This aligns with the well established increased importance of WRN in MSI-H tumors. However, compared to the amplitude of changes observed for PARP in homologous recombination-deficient (HR-deficient) ovarian cancer, the significance of this WRN signal is rather small.

In HR-deficient ovarian tumors, the frequency of PARP gene loss also decreases by half, but there is also a notable increase in PARP gene amplification (from 0% to 8%). The co-occurrence of reduced gene loss and gene amplification highlights the critical role of PARP in this cohort, supporting its attractiveness as a therapeutic target.

This contrast suggests that although WRN dependency is elevated in MSI-H colorectal cancers, its relevance as a drug target may be less pronounced than that of PARP in HR-deficient ovarian cancer.

Tumors were classified as MSI-High (MSI-H) based on bi-allelic loss of key mismatch repair genes such as MSH2, MSH6, MLH1, or PMS2. For HR-deficiency, classification was done using the HRD-score, calculated as the sum of genomic scar markers, including loss of heterozygosity (LOH), telomeric allelic imbalance (TAI), and large-scale state transitions (LST) (14, 15, 16).

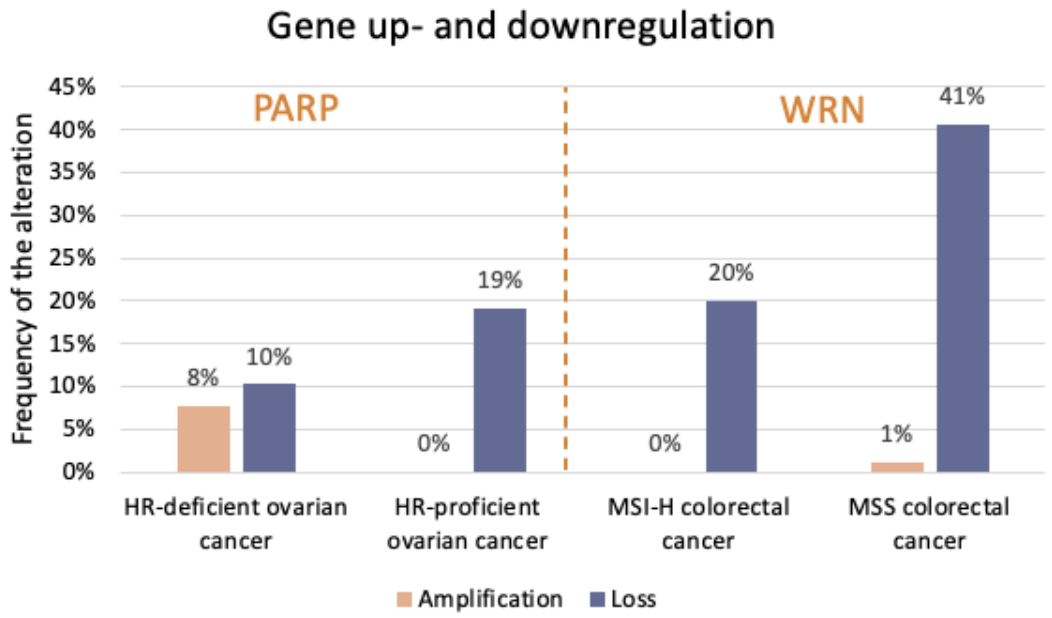


Figure 5. *Analysis of WRN gene alterations in colorectal cancer stratified by microsatellite instability (MSI) status (MSI-H - MSI High; MSS - MS Stable), compared with PARP alterations in homologous recombination (HR) deficient and proficient ovarian cancer.*

Validation experiments

The link between MSI-H tumors and WRN dependency has been extensively documented in the literature and validated experimentally in multiple cell lines. However, consistent with Gordion analysis, this dependency is too weak to lead to high clinical response rates, as confirmed by WRNi clinical trials (e.g., NCT06004245).

2. Increased essentiality of WRN in MYC-amplified ovarian patients

The Gordion algorithm evaluated the impact of various genomic alterations on WRN dependency in ovarian cancer. Common alterations, such as TP53 mutations and HR-deficiency events like BRCA1 or BRCA2 inactivation, only weakly increased WRN essentiality (17). An interesting signal emerged in MYC-amplified tumors (42% of ovarian tumors), which correlated with the removal of gene copies of WRN. That observation might lead to a therapeutic opportunity:

MYC amplification frequently occurs alongside whole genome duplication (WGD), leading to the presence of four copies of every gene in the genome (Detailed information regarding WGD is presented in the supplementary materials). MYC amplification typically arises through isochromosome 8 formation, which duplicates the long arm (8q) while deleting the

short arm (8p) containing the WRN locus (18). Interestingly, in samples after whole genome duplication, MYC is so frequently amplified via isochromosome 8 formation that from the initial four WRN gene copies, only a single copy of WRN remains. This will invariably lead to reduced WRN expression, which could lead to higher sensitivity of these ovarian cancer patients to WRNi. This could provide a therapeutic window that could deliver a high clinical response rate despite WRN not showing extremely strong essentiality in tumor data.

Market size for MYC-amplified ovarian cancer

The estimated annual new cases of ovarian cancer in the US are about 20,890. Approximately 42% of these new cases exhibit MYC amplification, which corresponds to around 8,774 new patients with MYC-amplified ovarian cancer each year.

Validation experiments

To confirm that WRN inhibition is more effective in MYC-amplified ovarian tumors, we propose a confirmatory experiment using engineered cell line models. This experiment could be conducted by our partner, VUS Genetics, a CRO with extensive expertise in gene editing and functional assays.

An isogenic MYC-amplified ovarian cancer cell line should be created: one retaining a single copy of the WRN gene, and a control line with all four copies intact. Should the WRN inhibitor demonstrate markedly increased cytotoxicity in the single-copy WRN model, this would provide a strong mechanism-of-action rationale for WRNi. Conversely, if differential sensitivity is not observed, we would recommend discontinuing the WRN inhibitor program, given prior failures in MSI-high colorectal cancer.

Ovarian cancer patient sub-population likely to respond to WRN inhibition.

We investigated the effect of the genetic mutations affecting WRN essentiality in ovarian tumors. Two strategies were deployed to identify genes increasing/decreasing the essentiality of WRN:

i) We search for the genetic context where **WRN amplifications are overrepresented, suggesting tumor dependence on high levels of WRN protein**.

ii) We search for the genetic context where **the loss of WRN is underrepresented, suggesting susceptibility to WRN inhibition**.

We explored a scenario where a mutation in a different gene leads to increased WRN essentiality, resulting in elevated rates of WRN amplifications or lack of the WRN losses. Alternatively, in cells susceptible to WRN inhibition, the loss of even a single copy should be

poorly tolerated. We analyzed which alterations are good predictors of the absence of WRN mono-allelic losses. We hypothesize that these mutated genes could serve as biomarkers for a cohort in which WRN inhibition might be more effective. To identify such genes and mutations, we performed a comprehensive analysis of the ovarian cancer cohort, quantifying WRN amplification or loss depletion frequencies in tumors harboring specific gene mutations

We considered three types of gene alterations for co-mutated genes (biomarker candidates):
● First, a bi-allelic loss of a gene, where all copies of a gene are lost. Genes that lose both copies in at least 1% of tumors were considered.
● Second, some genes are known to be haploinsufficient and a mutation in one copy already confers a beneficial effect on tumor growth (9,10). For a predefined list of such genes, we considered tumors with mono-allelic loss of such genes. The list was enriched in genes involved in DNA repair processes. Gordion is using a proprietary list of haplosupressors (details in the Supplementary Materials).
● Third, we used gene activations, defined as genes with gain-of-function mutations or amplifications.

To test the relationship of a potential biomarker with the rate of WRN amplification or loss depletion, for each biomarker, we divided the tumors into two groups: biomarker-mutated (e.g., bi- or mono-allelic loss (combined SNV, SV, indels, deletions), or amplified) and non-mutated. A logistic regression model was built to predict WRN amplification or loss depletion based on the biomarker status and a confounding variable - the overall rate of WRN amplification or loss depletion in the tumor. P-values and odds ratios (OR) were used to prioritize the candidate biomarkers. For an optimal biomarker, a high OR, a low p-value, and a large cohort size are expected.

	Gene symbol	Gene alteration	Chr. arm	OR [CI]	p-value	Cohort size
RESPONSE BIOMARKERS	MDM2	amplified	12q	17.62 [0.86-359.2]	0.06	4%
	SMARCA4	mono-allelic loss	19p	2.55 [0.95-6.84]	0.06	9%
	BRCA2	mono-allelic loss	13q	13.33 [0.56-317.5]	0.1	20%
	CCND1	amplified	11q	8.43 [0.48-149.0]	0.2	6%
	PTEN	mono-allelic loss	10q	1.77 [0.78-4.0]	0.2	12%
	ARID1A	mono-allelic loss	1p	1.56 [0.81-3.0]	0.2	20%
	RSF1	amplified	11q	6.7 [0.38-118.06]	0.2	9%

Table 1. *A list of genes positively associated with **WRN** (only genes with OR > 1.5; p-value < 0.2 are shown). OR - odds ratio; CI - confidence interval; Cohort size is the percentage of tumors with a given gene alteration present in **ovarian patient samples**.*

The most promising biomarker candidate identified is MDM2 amplification (OR = 17.6, p = 0.06). This observation is consistent with published reports (11) indicating that MDM2 acts as an E3 ubiquitin ligase for the WRN protein, promoting its ubiquitination and leading to downregulation of WRN levels. Consequently, tumors harboring MDM2 amplification may be particularly susceptible to WRN inhibition, potentially enhancing therapeutic efficacy in this subset of patients. MDM2 is also a well-established negative regulator of P53, but the present analysis did not reveal a significant association between WRN and TP53 status.

Additional biomarker candidates include genes encoding components of the SWI/SNF chromatin remodeling complex, specifically within the cBAF submodule. Among these, SMARCA4 (OR = 2.55, p = 0.06), ARID1A (OR = 1.56, p = 0.2), and RSF1 (OR = 6.7, p = 0.2) emerged as potential biomarkers; however, none of these associations reached statistical significance.

BRCA2 loss was also identified as a putative biomarker (OR = 13.33, p = 0.1), though—similarly—without achieving statistical significance. Furthermore, associations were observed for CCND1 amplification and PTEN loss, both recognized as key cell cycle regulators. Importantly, none of the candidate biomarkers demonstrated a statistically significant relationship with WRN status at the p < 0.05 threshold.

Biomarkers demonstrating a strong inverse association with WRN (OR < 0.15) were designated as candidate resistance markers. In the presence of these alterations, WRN is more often lost than expected, leading us to hypothesize that WRN might not be particularly important for the survival of these tumors.

RESISTANCE BIOMARKERS	Gene symbol	Gene alteration	Chr. arm	OR [CI]	p-value	Cohort size
	KMT2C	mono-allelic loss	7q	0.1 [0.02-0.48]	0.004	6%

Table 2. *A list of genes negatively associated with **WRN** (OR < 0.15; p-value < 0.2) - potential resistance biomarkers. OR - odds ratio; CI - confidence interval; Cohort size is the percentage of tumors with a given gene alteration present in **ovarian patient samples**.*

Only KMT2C met the criteria as a relevant biomarker (OR = 0.1, p = 0.004). Full results are available upon request.

Conclusions

Gordion analysis revealed that WRN is likely a “false positive” drug target that is unlikely to deliver clinical benefits when inhibited. This was revealed by the analysis of mutations affecting the WRN gene, as well as the analysis of all other 20,000 genes, including potential compensatory pathways. Importantly, the analysis of compensatory pathways explains why CRISPR data provided such a strong signal, while in the clinic, very little benefit was delivered by WRNi. Furthermore, the limitation of publicly available tools and the superiority of Gordion’s algorithms revealed a high level of biallelic losses of WRN. This later observation itself should put an assumption that WRN is a good drug target into question.

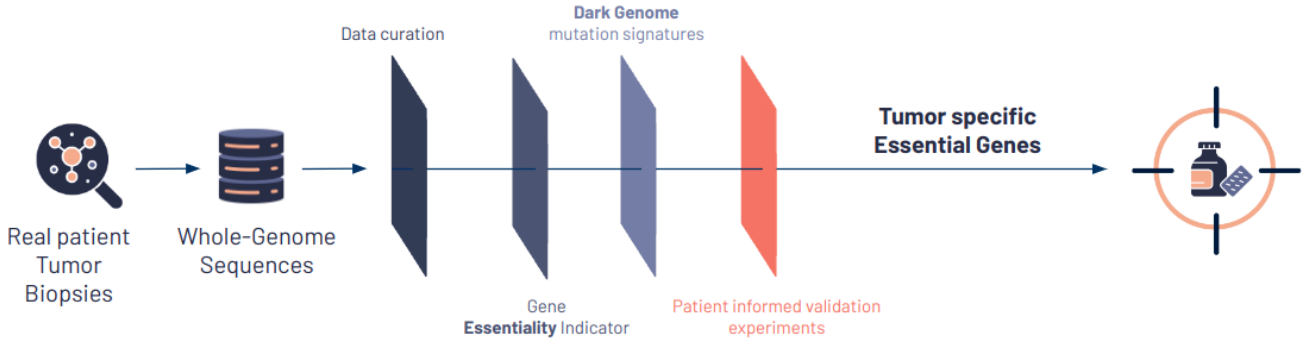
In order to identify a slightest opportunity for commercial success, a comprehensive pan-cancer dependency profiling was performed, revealing that ovarian and colorectal cancers are the tumor types likely dependent on WRN function. Further analysis of these cohorts identified subgroups with modestly increased WRN dependency: colorectal tumors with mismatch repair defects, MSI-H, and MYC-amplified ovarian cancers. Despite these findings, the observed dependency increase is considerably weaker (an order of magnitude) than that for established targets such as PARP inhibitors. Moreover, we identified several biomarkers that may enhance response rates to WRN inhibition, including MDM2 amplification, defects in chromatin remodeling processes (cBAF complex), and dysregulation of cell cycle control (PTEN loss or CCND1 amplification). BRCA2 loss was also identified as a potential response biomarker. However, associations with these biomarkers were relatively weak, below the threshold of statistical significance.

Given the comparatively weak signal from tumor genetic dependency analyses, the Gordion team concluded that inhibition of WRN is unlikely to substantially impact tumor growth. Consequently, the recommendation is to discontinue further investments into WRN inhibitor development. Potential drug combinations and resistance mechanisms were not assessed due to the low signal for WRN as a standalone target.

SUPPLEMENTARY INFORMATION

The Gordion Platform - overview

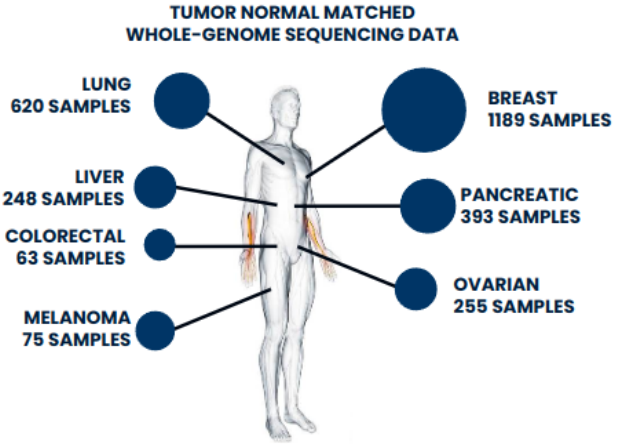
The Gordion Platform introduces an innovative method for evaluating oncology drug programs. By using the patient's whole genome data and harnessing the power of Machine Learning on the Dark Genome, it identifies essential genes specific to tumor types.



Supplementary Figure 1. Gordion Platform overview.

The Data

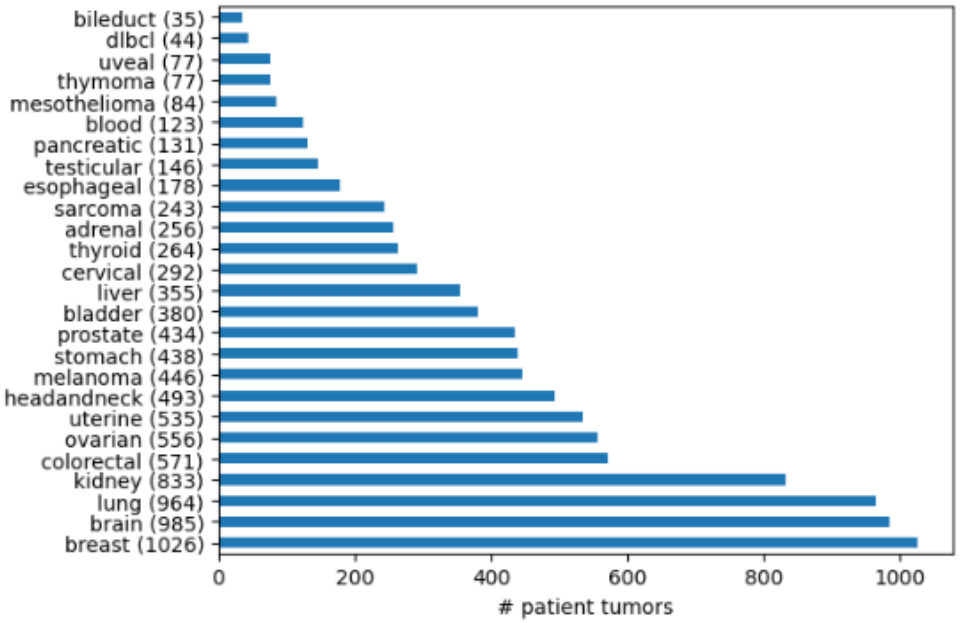
Gordion's platform is built on top of thousands of carefully selected and uniformly processed patient tumor whole genomes. These span 7 cancer types: breast, lung, ovarian, liver, pancreatic, colorectal, and melanoma, supported by 9,996 whole-genome CNV profiles from 26 tumor types.



Supplementary Figure 2. Tumor types and quantity of samples in the WGS Database.

- WGS Database - The 2,607 tumor-normal matched whole genome datasets are split between 7 cancer types (breast, ovarian, pancreatic, lung, colorectal, melanoma, and liver). All have been sequenced on the Illumina platform using paired-end reads with an average depth of 30-60x (tumors) and 25-60x (normal tissue). Samples have fully characterized genetic mutation profiles, including single-nucleotide variants, in-dels, copy number alterations, and structural changes.

- CNA+WES Database - whole-genome CNV combined with whole-exome sequencing data profiles for 9,966 tumor genomes across 26 cancer types. The raw data were originally produced by The Cancer Genome Atlas (TCGA) and were processed with the ASCAT3 tool.



Supplementary Figure 3. Tumor types and quantity of samples in the 10K CNA Database.

Data pre-processing and quality control

To build a large collection of tumor genomes across different cancer types, we aggregated datasets from multiple sources. These data sources have different application processes, as well as data storage/retrieval mechanisms. All three major repositories of genomic data have proprietary tools for accessing it: gdc tool is used to download data from TCGA; score from ICGC; and pyEGA3 for downloading data deposited in EGA.

Due to specific research problems addressed by academic projects, data generated by diverse research groups, also as part of one consortium, can substantially differ in the study design, data processing methods, and format of the released data. Aggregation of such datasets requires extensive quality checks and uniform re-processing and typically results in

excluding >10% of patient records. Adding a new dataset from a public repository to Gordion's database has often required solving project- or sample-specific problems with the downloaded raw data.

The typical input data for the WGS processing pipeline is the raw output of the DNA sequencer - short sequences (reads) in the FASTQ format. The WGS data deposited into repositories is often aligned reads (BAM or CRAM format). The alignments can be prepared in different ways, using different reference genomes, and this has a critical effect on the downstream processing.

To maintain uniformity of the results on Gordion's Platform, we re-align all data that is released in BAM or CRAM formats. Processing such datasets starts with the extraction of raw reads and base quality values (in the form of two FASTQ files) from the original BAM or CRAM files, and is described in the section below

Extraction of reads from alignment files
First, the BAM or CRAM file is checked against invalid reads that may break the downstream processing. In the case of CRAM input, the exact reference genome file that was used for the alignment must be identified to facilitate processing. The filtered alignment files are reindexed and subjected to read extraction using Bazam. Using this tool allowed us to avoid costly read-sorting (input alignments are sorted by genomic coordinates, and FASTQ files require name-sorted reads) and a very large intermediate file, saving both compute and storage costs.

As a result of Bazam processing, the reads from the BAM/CRAM file are split into a pair of gzipped FASTQ files containing read "mates". Next, the reads in the FASTQ files are trimmed to remove low-quality and adapter sequences. Samples with coverage depth greater than 62x undergo random downsampling to ~60x. A fixed random seed is used to facilitate the reproducibility of this process. Finally, data quality is analyzed and saved in QC reports.

Purpose(s)	Software	Reference
Alignment manipulation	Samtools 1.12	Danecek et al., 2021
Alignment validation		
Read indexing		
Read extraction from coordinate-sorted BAM	Bazam 1.0.1	Sadedin and Oshlack, 2019
Read quality trimming	Fastp 0.20.1	Chen et al., 2018
Read subsampling	Seqtk 1.3-r106	Github link

FASTQ quality control	FastQC v0.11.9	Andrews, 2010
QC report aggregation	MulitQC v1.10.1	Ewels et al., 2016

Supplementary Table 1. Tools used for WGS data pre-processing and quality control.

Sample eligibility criteria

To maintain a high-quality level of the data on the platform, we require that the added samples meet certain criteria. Depending on the data source and available information, some of them can help in the identification of non-eligible samples up-front, while others lead to sample exclusion after pre-processing (e.g., low read count) or only after processing with our WGS analysis pipeline (e.g, uniformity of coverage). Examples of the criteria required include:

- Fresh-frozen material
- Paired-end Illumina sequencing with at least 75bp long reads
- Mean tumor genome depth after deduplication >= 30x
- Mean normal genome depth after deduplication >= 25x
- Percent of the normal genome covered by at least 20x >= 80%
- Percentage of mapped reads >90%
- Tumor purity >= 20%
- germline variants between 4M and 5.5M
- somatic variants within expected ranges for the tumor type

The WGS analysis pipeline

The whole genome sequencing (WGS) pipeline maps the short reads to the reference genome and detects and genotypes all kinds of genetic variants - substitutions, short indels, copy number alterations, and other structural variants, both somatic and germline. All variants are quality-filtered and annotated. Quality control (QC) metrics are generated at all stages of the processing.

The WGS pipeline allows for processing sequencing data from pairs of matched tumor and normal samples. It also implements an experimental tumor-only mode. It is optimized for data from Illumina platforms and has been validated in the analysis of >6000 genomes with a mean genome coverage of 30-60x.

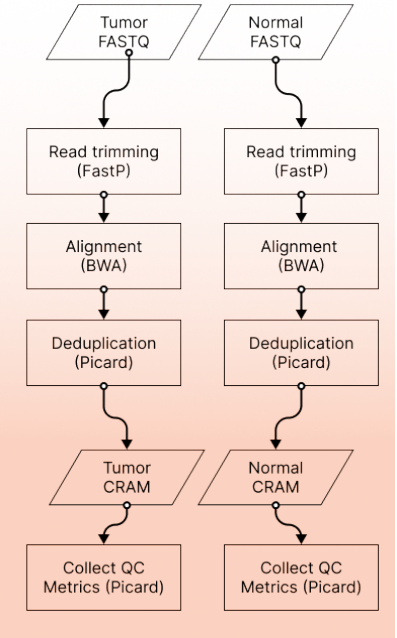
Description of the pipeline is divided into two stages:

- alignment stage that takes as input reads in FASTQ format, maps them to the reference genome (GRCh37), and post-processes the alignment file, producing a CRAM file ready for subsequent analysis
- variant calling and annotation stage, which uses a matched pair of produced CRAM files and generates a set of annotated variant files (VCF)

Alignment

Reads, provided in FASTQ files, are optionally trimmed and mapped to the GRCh37.d5 reference genome by the de facto standard in such applications BWA-MEM algorithm. The resulting alignment is sorted by genomic coordinates. Duplicate reads are marked using

Picard, and the output is saved to a CRAM file and indexed. After the CRAM has been created, coverage statistics are calculated, and a QC analysis is performed using Picard's CollectMultipleMetrics function. The QC statistics are saved to a report file in a PDF format, and in tab-delimited files for machine processing, e.g,. flagging low-quality alignments.



Supplementary Figure 4. Alignment workflow.

Purpose(s)	Software	Reference
Read quality trimming	Fastp 0.20.1	Chen et al., 2018
Read mapping		Li, 2013
Alignment indexing		Danecek et al., 2021
Read sorting Duplicate marking Alignment QC	Picard 2.22.4	
Depth of coverage	mosdepth 0.3.2	Pedersen and Quinlan, 2018

Supplementary Table 2. Tools used for the alignment.

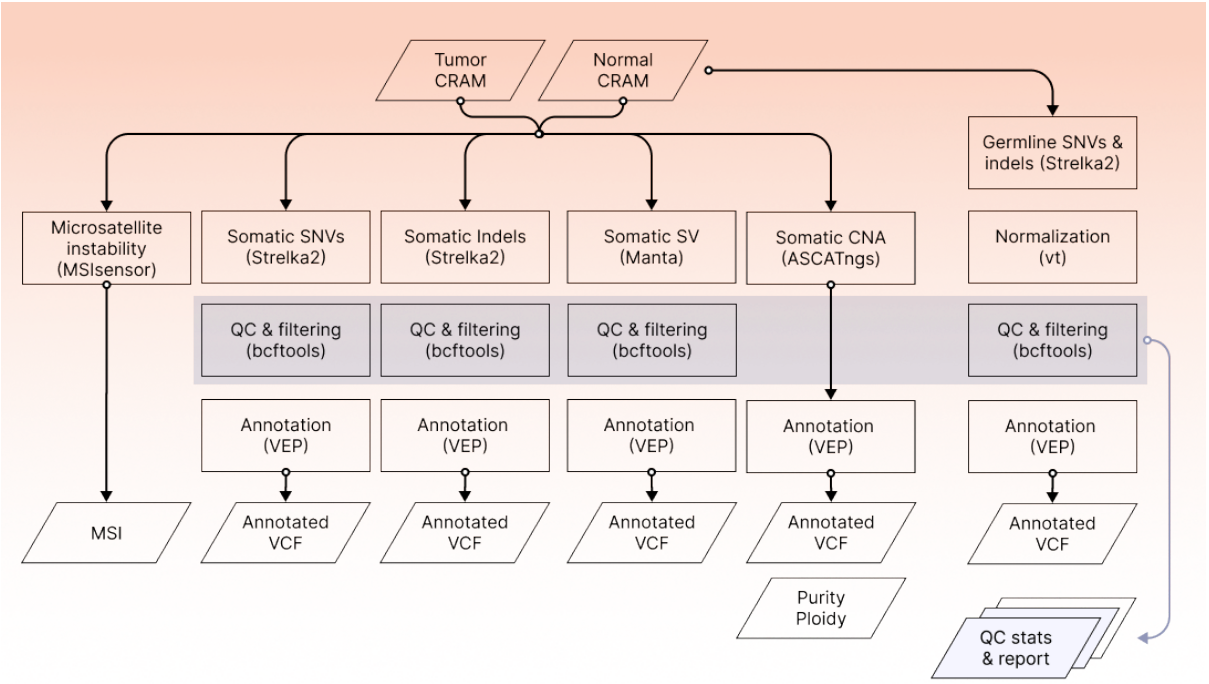
Variant calling

Our variant calling pipeline can work in either of two modes: matched tumor-normal analysis or tumor-only analysis. The latter has been used experimentally only and all data in our platform are based on a tumor-normal matched workflow.

Matched tumor-normal analysis

The matched analysis uses as input a pair of tumor and normal alignment files (BAM or CRAM) generated by the alignment workflow. By contrasting the tumor alignment with its matched normal alignment as a reference, individual algorithms detect changes in the tumor DNA (somatic variation). Five categories of variation - microsatellite instability (MSI), substitutions (SNVs), short indels, structural variants (SVs), and copy-number alterations (CNA) - are detected independently of each other by dedicated tools. Where applicable a quality control and filtering step (low-quality variants are removed) is performed. All variants passing quality criteria are subject to annotation, with (among others) Clinvar, population allele frequencies (1000 Genomes, ExAC, gnomAD, and two in-house databases of cancer patients), and various pathogenicity scores (e.g., SIFT, PolyPhen, DANN).

The pipeline detects and annotates germline variation as well. In the process, only the germline alignment file is used. Since germline substitutions & indels frequently contain multiallelic variants, for the purpose of annotation, these are decomposed and normalized. Next, similarly, as for somatic variant detection, the variants are filtered and annotated.



Supplementary Figure 5. Matched tumor-normal analysis workflow.

Individual algorithms used in the pipeline are implemented in open-source, commercially available software. They represent state-of-the-art methods used in projects published in

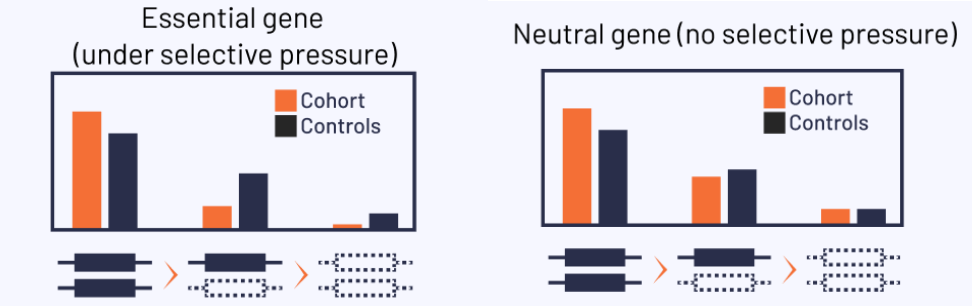
top-tier journals (e.g., PCAWG). Two of the tools (i.e., Strelka and Manta) were co-developed by Illumina Inc. and have laid the foundation for the Illumina Dragen Platform. The detailed list of tools is presented in the table below.

Purpose(s)	Software	Reference
Detection of: • Somatic SNV/indels • Germline SNV/indels	Strelka 2.9.10	Kim et al., 2018
Detection of: • CNA segments • Purity • Ploidy	ASCAT v3.0.0	Van Loo et al., 2010
Detection of: • Somatic SV • Germline SV	Manta 1.6.0	Chen et al., 2016647377/
Microsatelite-instability	MSI sensor 0.6	Niu et al., 2014
Variant manipulation		
Variant QC & filtering	BCFtools 1.12	Danecek et al., 2021
Variant annotation	VEP XXX.X	McLaren et al., 2016
QC report aggregation	MulitQC	

Supplementary Table 3. Tools used for the matched tumor-normal analysis workflow.

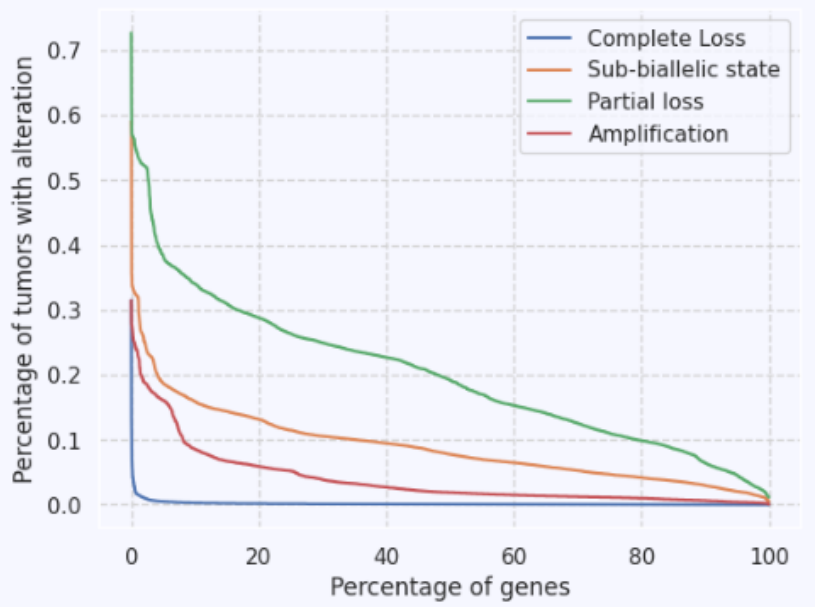
Partial loss

Essential genes are protected by the cell during tumor evolution. Looking at the rates of gene losses and amplifications, we can distinguish genes that are under positive selective pressure (essential genes; rarely lost/mutated) from the rest. Partial loss (pLOSS) of a gene is a concept adopted at Gordion to capture pressure on tumor cells to retain functional copies of essential genes. It allows us to measure selective pressure in all genes, including those that very rarely lose all copies.



Supplementary Figure 6. Selective pressure on essential and neutral genes.

Very few genes are lost completely (bi-allelic loss, also known as complete loss (cLOSS)) in a sufficient number of tumors to facilitate robust statistical analysis. This means that for the majority of genes, one would have to collect an infeasible number of tumor genomes to perform an analysis. However, partial losses (one or more copies) are much more frequent, with 80% of genes having lost at least a single copy of the gene in 10% of all the tumors in our database.



Supplementary Figure 7. Relationship between gene alterations and the number of affected tumors. For 80% of the protein-coding genes, at least 10% of Gordion's cohort (N=2607) has a partial loss. 40% of the genes are in the sub-biallelic state in at least 10% of the tumors. 20% of the genes have an amplification in at least 6% of the tumors, and only 19 genes (~0.5%) are completely lost in more than 10% of the tumors.

At Gordion, we have defined the concept of partial loss, also called single copy loss, as a state of the gene where the number of functional copies of the gene is below the average ploidy of the tumor. Gene loss leads to lower expression and reduced availability of the protein, which, in the case of essential genes, induces stress and lower cell fitness. In highly mutating tumor cells, it also exposes the cells to lethal complete gene loss, so it is eliminated by the tumor evolution. Depletion of partial gene losses (and enrichment of amplifications) is thus a marker of essentiality; one that can be measured for the majority of genes in cohorts of >100 patients.

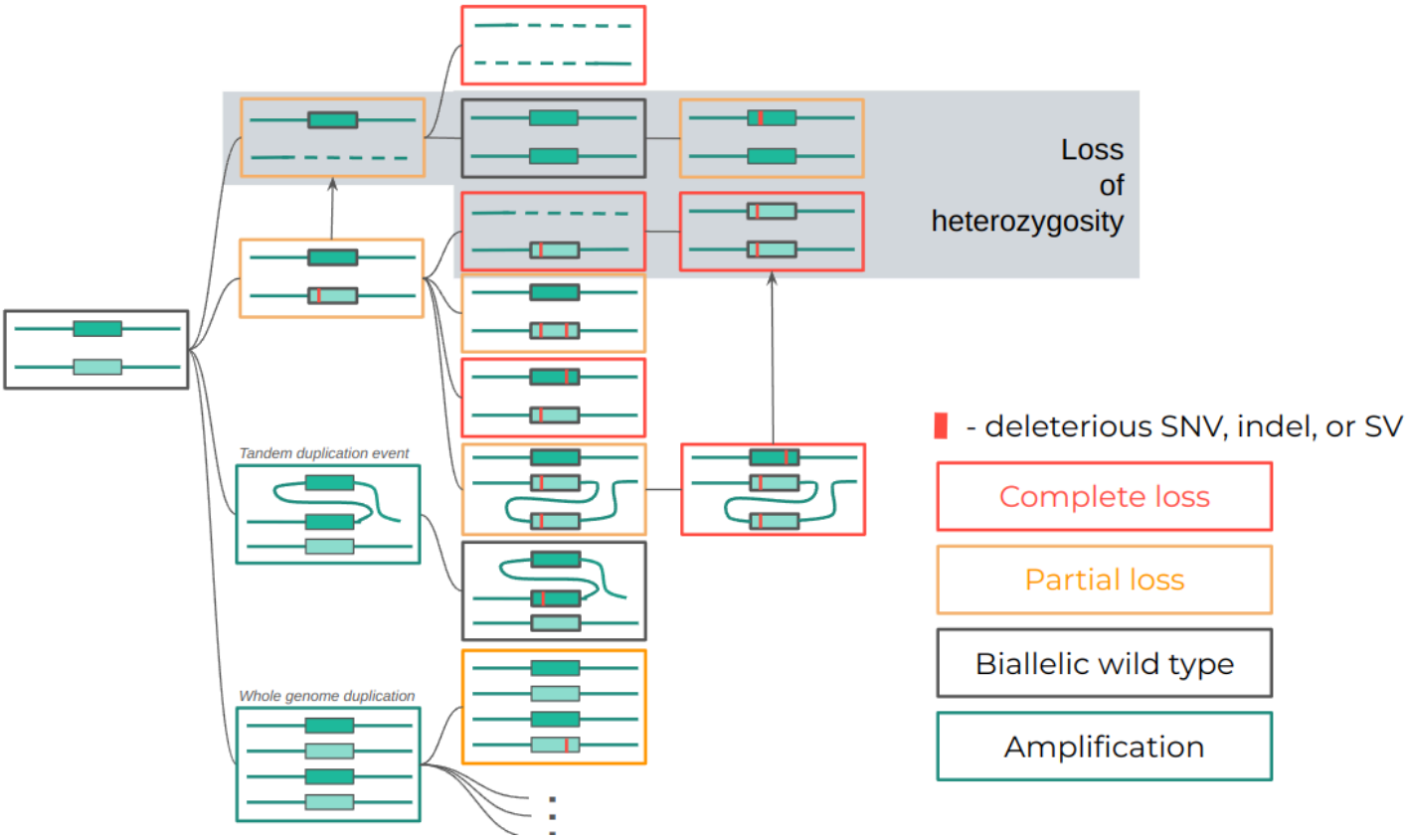
On the DNA level, the function of a gene can be lost in multiple ways, for example:

- the presence of a protein-altering single-nucleotide variant (SNV) or a short indel,

- by large structural variants (SVs), affecting the gene sequence, promoters, or enhancers,
- by copy number alterations (CNA), including, e.g., copy-neutral loss-of-heterozygosity.

Gordion is also analyzing single-nucleotide variants and their effect on gene function. However, besides deciding whether a variant itself affects a protein's function, in tumors, one needs to take into consideration additional biological and technical factors:

- one or more alleles can be affected - more than half of tumors undergo a genome duplication event
- two mutations in one gene can be in cis (alter the same copy; one copy is lost) or in trans (each modifies a different copy; at least 2 copies are lost)
- germline and somatic alterations need to be evaluated jointly because variant detection tools subtract the germline variation as background from their results
- somatic mutations may affect the main clone (the majority of tumor cells) or be subclonal (only a fraction of the tumor cells carry the variant)
- sample purity - the fraction of tumor cells in the sequenced sample can vary between 20 and 100%



Supplementary Figure 8. Examples of loss events and selection of mutational trajectories leading to them.

Overall, the Loss of Function outcome is a product of the events mentioned above. Therefore, characterizing the number of functional gene copies in a tumor genome has been a tremendously complex, laborious, and error-prone process involving field experts, from geneticists to bioinformaticians. The analysis requires manual integration of the results from several variant callers, careful and time-consuming evaluation of each, and often dealing with contradictory evidence when comparing the event's overall impact.

We have solved this by creating a proprietary 'FunctionalCopyProfiler'. Utilizing in-house expertise gained from profiling hundreds of tumor genomes, we have developed a tool that combines results across the different variant callers we use. For both somatic and germline genomes, FunctionalCopyProfiler processes small variants (SNVs and indels detected by Strelka2), large structural variations (deletions, tandem-duplication, inversions, translocations, insertions; detected by Manta), and copy-number changes (ASCAT tool), all supplemented with rich functional annotations, and the tumor's ploidy and purity. FunctionalCopyProfiler can infer the number of functional and lost copies of all protein-coding genes in the cancer genome using a broad set of customizable expert rules.

Not all pathogenic mutations lead to loss of function

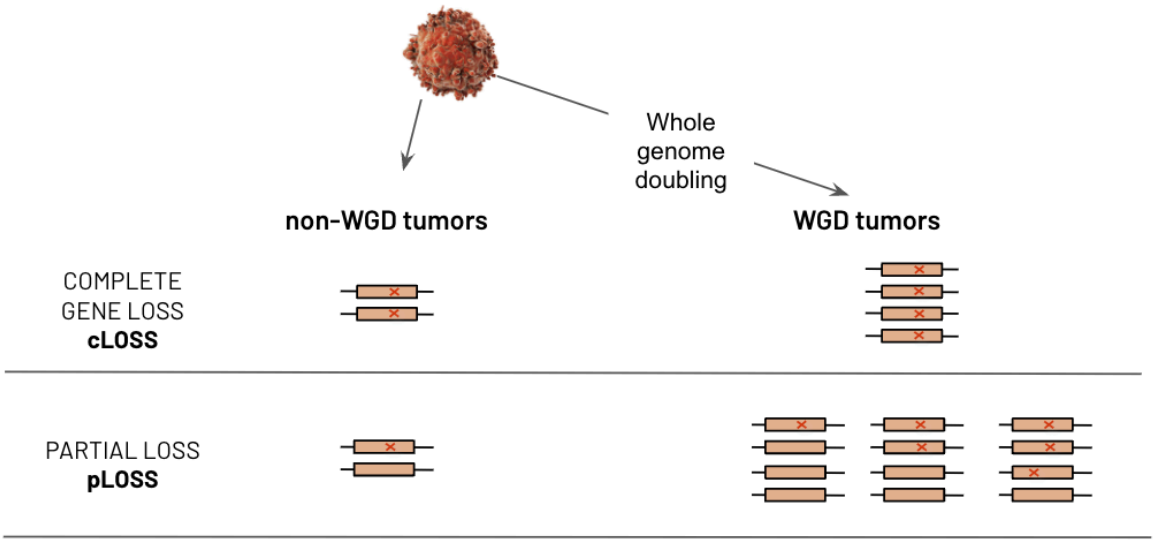
A typical understanding of pathogenic mutation is that it ablates protein function. For the majority of the genes and phenotypes, this is correct; however, in many cases, the phenotype is caused by a gene's activation (KRAS, BRAF) or taking on a novel function (NPM1), Sometimes, the structural variants can lead to fusion genes, which encode hybrid proteins, affecting the cell's metabolism and fuelling the oncogenesis. There are also proteins where mutations, all classified in mutation databases as pathogenic, lead to either loss of function or gain of function, therefore eluding the commonly accepted division into oncogenes and tumor suppressor genes (TP53, DNMT3A). Such situations must be taken into account when classifying a gene as lost.

Mutations classified as pathogenic and acting via a mechanism different from protein function loss are not considered to lead to (partial) gene loss. For critical oncogenes (e.g., KRAS) and tumor suppressor genes (eg, TP53), the Gordion team created and maintains a manually curated list of mutations that are classified as loss-of-function (LOF), gain-of-function (GOF), and unknown. These classes of mutations are treated differently by the platform.

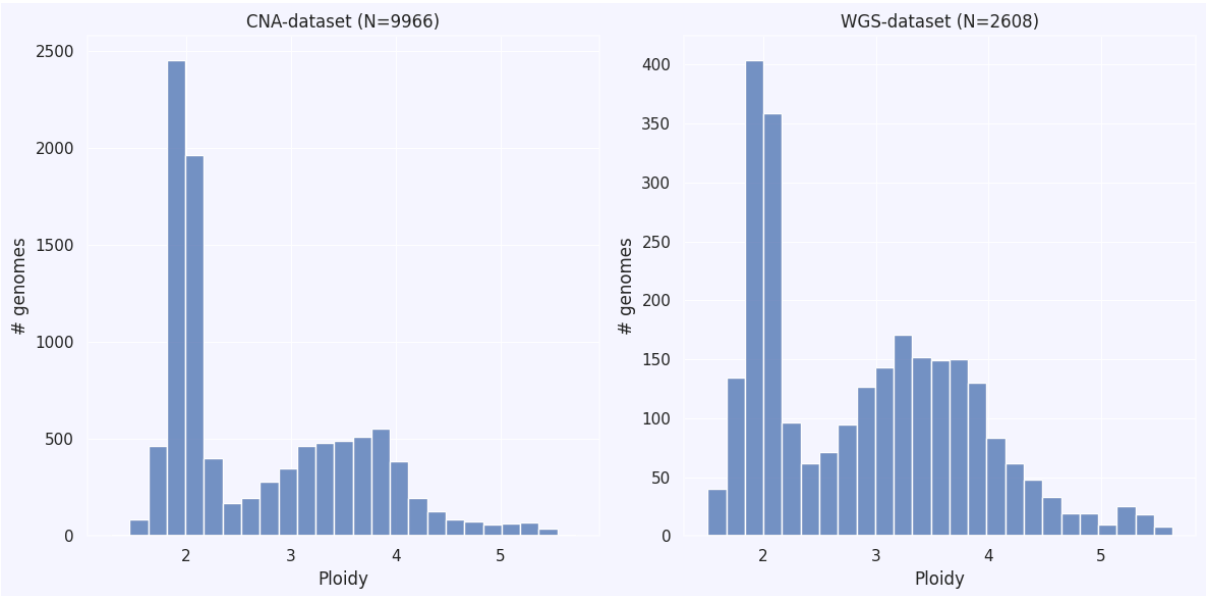
Whole Genome Duplication

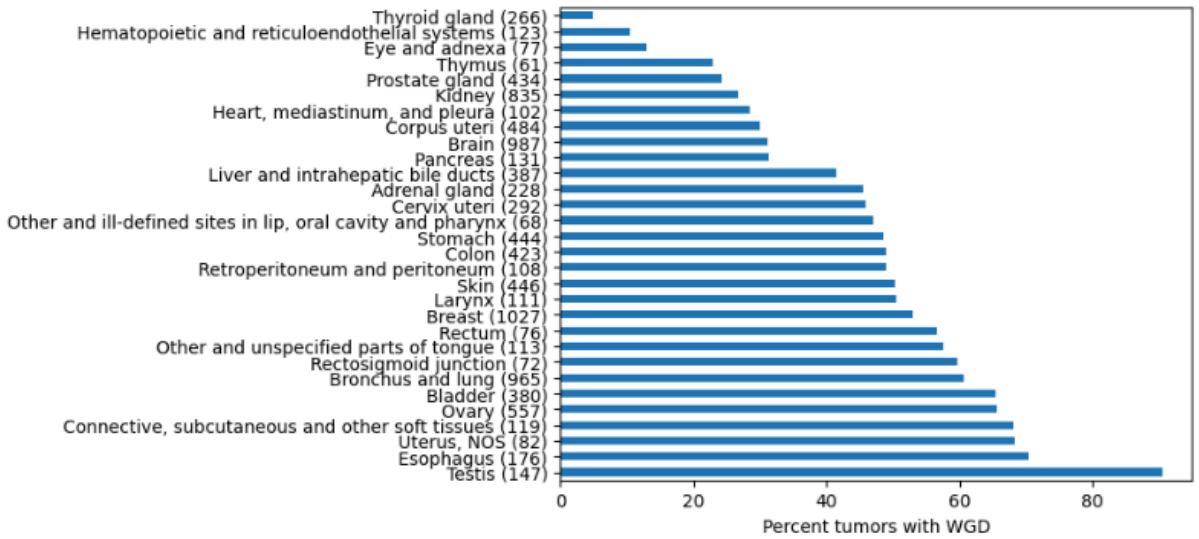
Whole Genome Duplication (WGD), also known as polyploidy, is a biological phenomenon wherein an organism's entire set of chromosomes is duplicated. This process results in multiple copies of the genome within a cell or organism. While WGD is relatively rare in normal human cells, it has been observed to play a crucial role in the context of cancer, with over 50% of solid tumors undergoing WGD events. WGD results in multiple copies of the genome within a cell (in our data, 41% of tumors are diploid, 57% have undergone WGD, and 2% are classified as above 4n). To address this challenge and properly quantify

deletions, we consider deletions that remove one copy from a diploid tumor sample and deletions that remove 1, 2, or 3 copies from samples with ploidy ≥4.



Supplementary Figure 8. Gordion utilizes different types of gene loss depending on tumor ploidy.





Supplementary Figure 9. Whole Genome Duplication in our data. The upper left panel shows a distribution of ploidy in the 10K CNA Database. Upper right in the WGS Database. The lower panel presents the percentage of tumors with WGD in different tumor types.

Understanding the implications of WGD becomes paramount when analyzing gene mutations in cancer for several reasons:

- **Tumor Evolution and Heterogeneity.** WGD can contribute to tumor evolution and heterogeneity. The presence of multiple copies of the genome allows for a higher degree of genetic variation. This heterogeneity poses challenges in predicting the behavior of the cancer and devising effective treatment strategies.

- **Impact on Mutational Landscape.** WGD influences the mutational landscape of a cancer cell. The increased genomic material provides more opportunities for mutations to arise, and certain mutations may be selectively advantageous, contributing to tumor progression and adaptability.

- **Therapeutic Implications.** The presence of WGD can affect the response to therapy. Tumor cells with duplicated genomes may exhibit altered sensitivity or resistance to certain treatments, making it crucial to consider WGD in the development of targeted therapies and personalized medicine approaches.

Accounting for WGD is essential for the accurate interpretation of genomic data. Failure to consider WGD may lead to misinterpretation of mutational profiles and hinder the identification of truly essential genes in the process of drug target ID. Gordion's partial loss, as well as gene amplifications, are calculated concerning the number of genome copies, taking into account genome doubling events.

Loss-of-function - our proprietary approach

Loss-of-function (LOF) of key genes is commonly used to stratify patients, for instance, to search for cohort-specific genomic biomarkers. Combining all types of variants is a very complex task, so the process typically involves only small variants (SNVs and indels). Even then, automation is far from trivial. The information in mutation databases used for classifying loss-of-function mutations is frequently incomplete or conflicting. Sequencing artifacts lead to false positive variant calls, which are difficult to assess without inspecting the raw data. Accurate classification of loss-of-function cannot be done without expert curation.
Our experts have performed manual curation of hundreds of tumors' whole genomes, going through mutation tables with extensive annotation, genome by genome. By cross-checking the variants' annotation with databases and publications, and inspecting the variants visually in the raw data, they interrogated several key cancer drivers and suppressor genes (e.g., TP53, KRAS, BRCA1, BRCA2), scrupulously annotating their mutational status in all samples. The extensive experience gained in this process allowed us to formulate a set of complex rules for the classification of variant combinations as loss-of-function events. Importantly, we have also specified rules that determine the lack of loss in cases that otherwise would result in false positive classifications. All these have been written down as an easily customizable set of "configurations" for FunctionalCopyProfiler.

The rules were then implemented directly in the tool to allow easier future changes both in variant detection algorithms, in variant annotations, and in the rules themselves. The initial version of the rules has been improved in several cycles of testing and benchmarking against a set of BRCA1 and BRCA2 losses in a group of over 800 breast cancer patients (created in the abovementioned manual profiling process). The current version yields 98.9% accuracy (96.2% sensitivity and 99.6% specificity) in classifying variants into three categories: complete-, partial-, and no-loss.

For each analyzed patient and gene, FunctionalCopyProfiler outputs the total number of copies and the number of functional copies, as well as auxiliary information on the mutations responsible for gene losses, and a degree of confidence in their classification. The statistics on the numbers of lost/functional copies aggregated over cohorts of patients are then used to measure the essentiality of genes.

GORDION AI MODEL

The Gordion AI stack integrates diverse analytical tools to evaluate multiple modalities of patient data. The model employs a suite of proprietary algorithms designed to detect novel molecular descriptors—unique patterns of mutations that may be associated with disease trajectory, tumor resistance, or specific molecular pathways.

Traditional analytical approaches face significant limitations when processing complex genomic patterns. The transformation of intricate descriptors into simplified features for conventional feature-based classifiers, or the reliance on domain-specific distance measures

for distance-based classifiers such as Similarity Forests, results in substantial information loss. To address this challenge, we developed innovative techniques, including Random Similarity Forests: Publication: Random Similarity Forests.

Prediction of expected gene copy number

The Gordion AI incorporates sophisticated tools to estimate expected gene copy number, a critical parameter for determining gene essentiality. While point mutations account for gene loss in fewer than 20% of cases, large-scale copy number variations (CNVs) are responsible for approximately 80% of gene losses. Depending on tumor type, specific chromosomal regions can be lost in 70-80% of tumors or amplified in up to 50%, providing compelling evidence of mutational pressure.

Large-scale CNVs play a fundamental role in determining gene essentiality. The extent of these alterations and their impact on gene dosage, focalization, and genomic context are key factors in distinguishing genes under strong selective pressure from those frequently lost or amplified as passenger events.

Comprehensive structural analysis beyond coding sequences

Our CNV analysis extends beyond standard approaches that focus primarily on coding sequences. We examine structural alterations across the entire genome, including large intergenic regions and intronic sequences, which can have significant regulatory consequences. Importantly, CNVs can influence enhancer activity, leading to dysregulated gene expression. A representative example is the identification of a specific amplification on chromosome 14q, which affects a super-enhancer associated with NKX2-1, a critical transcription factor involved in lung development and oncogenesis.

Driver gene identification from copy number alterations

Copy number alterations (CNAs) frequently affect large chromosomal fragments, often encompassing entire chromosomal arms or even complete chromosomes, making it challenging to pinpoint causal genes. By integrating mutational pressure across two modalities—CNAs and point mutations—we identify drivers of CNA events. Using this information and the observed copy numbers of all genes along the chromosome, we construct AI models that predict the expected gene copy number for every gene. Models are built separately for each chromosomal arm and tumor type. When applied to an analyzed cohort, the frequency of driver mutations is used to predict the expected copy number profile along chromosomes. Observed copy numbers are compared with expected rates to highlight genes and entire loci under positive selective pressure.

Detection of cohort-specific drivers

For point mutations and indels, the background mutation rate is known to correlate with sequence context and chromatin accessibility Supek F, et al. DNA Repair 2019; Polak P et

al. Nature 2015. We utilize tumor-type-specific and pan-cancer maps of expected mutation rates generated by AI models trained on hundreds of epigenetic profiles and nucleotide composition patterns.

These maps enable precise calculation of mutation excess and depletion in any given cohort, facilitating the identification of tumorigenesis drivers Sherman MA, et al. Nat Biotechnol. 2022. Genes that accumulate an excess of truncating mutations (e.g., stop-gain, frameshift, splice-site) are likely tumor suppressors negatively selected during tumor evolution (e.g., TP53, SMAD4)—tumor cells disable the activity of these genes. Activating (gain-of-function) mutations accumulate at specific positions in oncogenes (e.g., G12C substitution in KRAS). In contrast to truncating mutations in suppressor genes, activating mutations are typically missense, but can also be non-coding (e.g., TERT promoter) or result from gene fusion events.

Maps of expected rates of small variants enable precise attribution of mutational pressure to genes and other functional DNA elements (promoters, enhancers), proving highly effective in identifying drivers. However, for the majority of genes, background mutation rates are too low to robustly demonstrate depletion of mutations. In these cases, copy number changes, including partial losses and shallow amplifications, provide valuable complementary information.

AI model for biomarker identification

Susceptibility to inhibition of the gene of interest (GoI) is assessed by measuring enrichment in GoI amplifications (indicating likely dependence on elevated GoI protein levels) and GoI retention (depletion in GoI losses). To evaluate the relationship between biomarkers and GoI amplifications and deletions, tumors are stratified into thousands of categories based on biomarker mutation status (e.g., bi-allelic gene loss, mono-allelic gene loss, amplification, or activation) and compared to non-mutated controls.

For each biomarker-GoI alteration pair, a logistic regression model is constructed to predict GoI amplification (and GoI retention) based on two predictors: biomarker status and a confounding variable—the overall rate of amplifications (or partial losses) in the tumor. This analysis is performed both pan-cancer and within individual tissue types. P-values and odds ratios (ORs) are used to prioritize candidate biomarkers.

The technological details of the presented technology are part of Gordion IP and are not publicly available.

REFERENCES

1. Teotia S, Wang F, Mekonnen Z, et al. Roles of RECQ helicases in genome maintenance and cancer therapy. Genes Dev. 2016;30(15):165–173.

2. Oshima J, Sidorova JM, Monnat RJ Jr. Werner syndrome: clinical features, pathogenesis and potential therapeutic interventions. Mech Ageing Dev. 2016;159:95–108.

3. Hu W, Rolig AS, Levy S, et al. Small-molecule inhibition of WRN helicase induces replication stress and selective toxicity in cancer cells. Nat Chem Biol. 2017;13(2):184–193.

4. Chan EM, Shostak Y, Smith A, et al. WRN helicase is a synthetic lethal target in microsatellite unstable cancers. Nature. 2019;568:551–556.

5. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546–1558.

6. Nicolae C, Stein A, Engels H, et al. DIG: a deep learning framework for genome-wide modeling of somatic mutation rates and driver discovery. Nat Genet. 2022;54:1234–1242. doi:10.1038/s41588-022-01085-7.

7. The Cancer Genome Atlas Research Network. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013;45:1113–1120.

8. Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discov. 2012;2(5):401–404.

9. Helsten T, Elkin S, Arthur E, et al. The FGFR landscape in cancer: analysis of 4,853 tumors by next-generation sequencing. Clin Cancer Res. 2016;22(1):259–267.

10. Newman JA, Erve J, Tanizawa Y, et al. RecQ helicases in DNA repair and cancer targets. Essays Biochem. 2020;64:819–838.

11. Lieb S, Baumann P, Mueller S, et al. Werner syndrome helicase is a selective vulnerability of microsatellite instability-high cancer cells. eLife. 2019;8:e43333.

12. El-Brolosy MA, Stainier DYR. Genetic compensation: A phenomenon in search of mechanisms. Dev Cell. 2017;41(3):275–279.

13. Smits AH, et al. Biological plasticity rescues target activity in CRISPR knockouts. Nat Methods. 2019;16(11):1087–1093.

14. Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br J Cancer. 2012;107(10):1776–1782.

15. Telli ML, Timms KM, Reid J, et al. Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. Clin Cancer Res. 2016;22(15):3764–3773.

16. Popova T, Manié E, Rieunier G, et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. Cancer Res. 2012;72(21):5454–5462.

17. Miller RE, Bellini E, Ginsburg E, et al. Ovarian cancer therapy: homologous recombination deficiency as a predictive biomarker for response to PARP inhibitors. JCO Precis Oncol. 2022;6:e2200112.

18. Tuna M, Amos CI, Mills GB. Whole-chromosome arm acquired uniparental disomy in cancer development is a consequence of isochromosome formation. Neoplasia. 2022;25:9–17. doi:10.1016/j.neo.2021.12.009.