

Genome analysis

TiMEx: a waiting time model for mutually exclusive cancer alterations

Simona Constantinescu^{1,2}, Ewa Szczurek^{1,2}, Pejman Mohammadi^{1,2},
Jörg Rahnenführer³ and Niko Beerenwinkel^{1,2,*}

¹Department of Biosystems Science and Engineering, ETH Zürich, ²Swiss Institute of Bioinformatics, Basel 4058, Switzerland and ³Faculty of Statistics, Technische Universität Dortmund, Dortmund 44221, Germany

*To whom correspondence should be addressed.

Associate Editor: Jan Korbel

Received on March 19, 2015; revised on June 12, 2015; accepted on June 26, 2015

Abstract

Motivation: Despite recent technological advances in genomic sciences, our understanding of cancer progression and its driving genetic alterations remains incomplete.

Results: We introduce TiMEx, a generative probabilistic model for detecting patterns of various degrees of mutual exclusivity across genetic alterations, which can indicate pathways involved in cancer progression. TiMEx explicitly accounts for the temporal interplay between the waiting times to alterations and the observation time. In simulation studies, we show that our model outperforms previous methods for detecting mutual exclusivity. On large-scale biological datasets, TiMEx identifies gene groups with strong functional biological relevance, while also proposing new candidates for biological validation. TiMEx possesses several advantages over previous methods, including a novel generative probabilistic model of tumorigenesis, direct estimation of the probability of mutual exclusivity interaction, computational efficiency and high sensitivity in detecting gene groups involving low-frequency alterations.

Availability and implementation: TiMEx is available as a Bioconductor R package at www.bsse.ethz.ch/cbg/software/TiMEx.

Contact: niko.beerenwinkel@bsse.ethz.ch

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

Despite recent technological advances in genomic sciences, our understanding of cancer progression still faces fundamental challenges. To this end, new ways of interpreting the increasing amount of generated data are devised, aiming at finding biologically relevant patterns. An important example is the separation of genes into drivers, which have a selective advantage and significantly contribute to tumor progression and passengers, which are selectively neutral and can hitchhike along with fitter clones. Even if intuitive and routinely used, identifying drivers as recurrently altered genes (Sjöblom *et al.*, 2006) only explains tumorigenesis in a fraction of patients. Alternatively, the functional role of drivers can be assessed in the context of groups of genes, all possessing the same important function, commonly known as pathways. Once one of the group members is altered, the tumor gains a significant selective advantage.

The alteration of additional group members does not further increase the selective advantage of the tumor, rendering clones with a single alteration likely the most frequent. In this case, the group of genes displays a mutually exclusive alteration pattern.

Current approaches for detecting mutual exclusivity are either *de novo* (Ding *et al.*, 2008; Jerby-Aron *et al.*, 2014; Leiserson *et al.*, 2013; Miller *et al.*, 2011; Szczurek and Beerenwinkel, 2014; Vandin *et al.*, 2012; Yeang *et al.*, 2008) or based on biological interaction networks (Ciriello *et al.*, 2012). While highly informative, the current biological knowledge is incomplete, such that limiting the search space to known biological interactions significantly reduces the detection power. Straightforward pairwise statistical tests assessing whether the number of observed double mutants is lower than expected by chance have also been employed, followed by identifying groups as maximal cliques (Ciriello *et al.*, 2012; Yeang *et al.*,

2008). The Dendrix tool (Vandin *et al.*, 2012) performs a Markov chain Monte Carlo sampling for group structure search, followed by a permutation test for finding sets of genes with both high coverage and high exclusivity. Its limitation of finding the single main pathway per dataset was addressed by Multidendrix (Leiserson *et al.*, 2013), a follow-up tool which simultaneously identifies multiple driver pathways via an integer linear programming approach. Szczurek and Beerenwinkel (2014) propose muex, a statistical model for mutual exclusivity, where, however, the group members are required to have similar alteration frequencies. Finally, Jerby-Arnon *et al.* (2014) employ a data-driven approach for identifying genome-wide synthetic lethality interactions in cancer, which lead to mutual exclusivity. All existing approaches ignore the fact that the mutually exclusive patterns occur over time, during disease progression.

Here, we introduce TiMEx, a generative probabilistic model for the *de novo* detection of mutual exclusivity patterns of various degrees across carcinogenic alterations. We regard tumorigenesis as a dynamic process and base our model on the temporal interplay between the waiting times to alterations, characteristic for every gene and alteration type and the observation time. Under the assumption of rarity of events over short time intervals, TiMEx models the alteration process for each gene as a Poisson process. The waiting times are therefore modeled as exponentially distributed variables with specific rates, which correspond to the rates of evolution for each alteration. In our modeling framework, the temporal dynamics of each alteration process progresses from the onset of cancer, corresponding to the first alteration related to the growth of a malignant tumor, until the observation time, corresponding to the time of tumor biopsy. The observation time is regarded as a system failure time and is exponentially distributed with an unknown rate.

A perfectly mutually exclusive group is defined as a collection of genes in which, for every tumor sample, at most one gene is altered. Conversely, we assume that in a group showing no mutual exclusivity, each gene is altered conditionally independent, given the observation time. In a realistic biological setting, however, additional alterations may still provide a small selective advantage to the tumor, rather than none at all, which may lead to the fixation of a clone with more than one alteration, in a group of genes otherwise perfectly mutually exclusive. Thus, biologically, groups of genes display a continuous range of mutual exclusivity degrees. TiMEx quantifies these degrees exactly and assesses their significance using a likelihood ratio test. Our procedure for efficient search of mutually exclusive patterns in large datasets consists of three steps (Fig. 1). We first estimate mutual exclusivity between all possible gene pairs

in the dataset. Second, we select as candidates the gene groups in which the significance and degree of mutual exclusivity between each pair of members are high. Third, the candidate groups are statistically tested for mutual exclusivity.

In simulation studies, we show that TiMEx outperforms the permutation-based method previously introduced by Vandin *et al.* (2012) and the muex model (Szczurek and Beerenwinkel, 2014). Furthermore, we apply our procedure to four large TCGA studies, two glioblastoma datasets, ovarian (Cancer Genome Atlas Research Network, 2011) and breast (provisional) cancer datasets. On these datasets, we show that TiMEx identifies gene groups with stronger functional biological relevance than two previous methods, namely muex and Multidendrix, while also proposing many new candidates for biological validation. TiMEx does not impose any temporal assumptions on the set of biological samples it is applied on, and these samples are considered to be independent. Without requiring any previous biological knowledge, our procedure identifies mutually exclusive gene groups of any size and statistically tests and ranks them by their degree of mutual exclusivity. It possesses several advantages over previous methods, including the probabilistic modeling of tumorigenesis as a dynamic process, the novel and intuitive quantification of the degree of mutual exclusivity as a probability, high computational efficiency on large datasets, as well as high sensitivity in detecting low frequently altered genes.

2 Methods

2.1 Probabilistic model

We consider n genes indexed by $N = \{1, 2, \dots, n\}$, whose alteration statuses are represented by the vector of binary random variables $X = (X_1, X_2, \dots, X_n)$, recorded at observation time $T_{\text{obs}} \sim \text{Exp}(\lambda_{\text{obs}})$. The waiting times to alteration of the n genes are represented by the vector of random variables $T = (T_1, T_2, \dots, T_n)$, where $T_i \sim \text{Exp}(\lambda_i)$, for all $i \in N$. For a given tumor sample, we refer to an instantiation of X , namely (x_1, \dots, x_n) , as a genotype, where $x_i \in \{0, 1\}$, for all $i \in N$. Moreover, for any set of indices $K \subset N$ with cardinality $|K|$, we denote by g_K the genotype for which positions with indices in K are equal to 1 and positions with indices in $N \setminus K$ are equal to 0. The presence of an alteration event in a tumor sample signifies both its occurrence in one of the tumor cells and its fixation in the measured population, such that the alteration is observed at screening. Let \mathcal{D} denote M independent observations $(X^{(1)}, \dots, X^{(M)})$. Each $X^{(j)} = (X_1^{(j)}, X_2^{(j)}, \dots, X_n^{(j)})$ denotes the

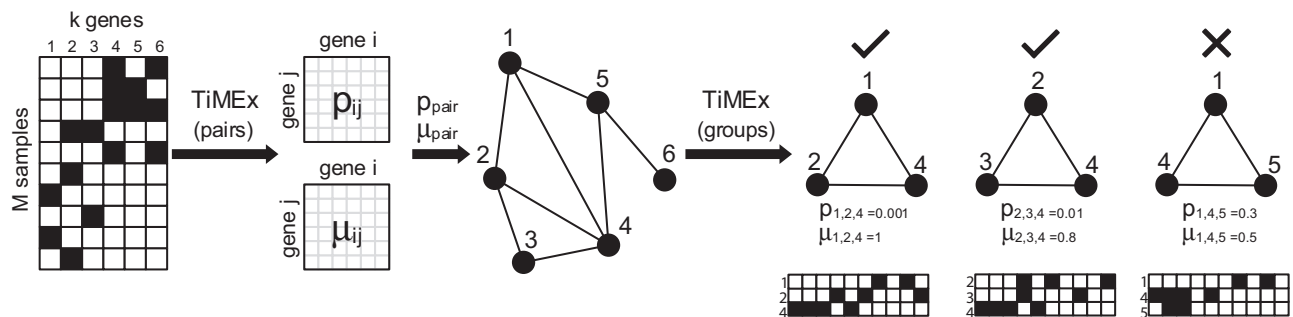


Fig. 1. Overview of the TiMEx multistep procedure for detecting mutually exclusive groups of alterations in a large dataset. First, from a binary alteration matrix consisting of M samples and k genes, the degree of mutual exclusivity μ_{ij} and the P value for testing $\mu_{ij} \neq 0$ against $\mu_{ij} = 0$ are estimated for all gene pairs (i, j) . Second, candidate groups are identified as maximal cliques of genes sharing a significant minimum degree of mutual exclusivity (satisfying the thresholds p_{pair} and μ_{pair} for each edge). Finally, the candidate groups are statistically tested for mutual exclusivity and the degree of mutual exclusivity corresponding to each group is estimated and tested for significance

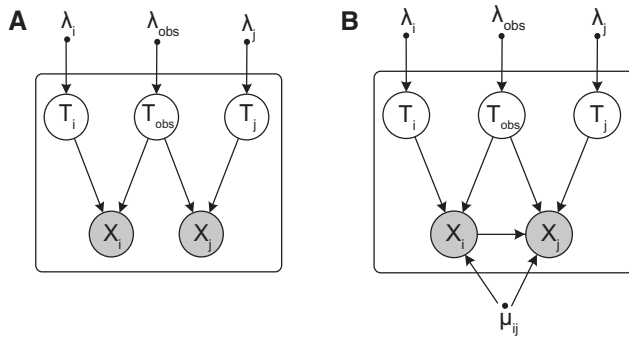


Fig. 2. Graphical representations of **(A)** null and **(B)** mutual exclusivity models for two genes, i and j . The observed variables are shaded in gray, while the hidden ones are not. The binary random variables X_i and X_j denote the alteration statuses of the two genes, T_i and T_j are their waiting times, and T_{obs} is the observation time. All variables are exponentially distributed with corresponding parameters λ . In the null model, X_i and X_j are conditionally independent given the observation time T_{obs} . In the mutual exclusivity model, X_i and X_j also depend on each other via the parameter μ_{ij} which represents their degree of mutual exclusivity

alteration statuses of the n genes in tumor sample j , and each $T^{(j)} = (T_1^{(j)}, T_2^{(j)}, \dots, T_n^{(j)})$ denotes their corresponding waiting times. The binary variables X_j are observed, while T_i and T_{obs} are hidden. We are interested in inferring the degree of mutual exclusivity among the group of n genes. To this end, we compute the likelihood of the data \mathcal{D} under the nested null and mutual exclusivity models introduced below. As the observations are independent, the likelihood of the data under any model θ is $L(\theta|\mathcal{D}) = \prod_{j=1}^M P(X^{(j)}|\theta)$.

2.1.1 Null model

The null model (Fig. 2A), parameterized by $\theta_{\text{Null}} = (\lambda_1, \lambda_2, \dots, \lambda_n, \lambda_{\text{obs}})$, assumes that alterations in the n genes are conditionally independent from each other, given the observation time T_{obs} . The condition for observing an alteration in a gene i is that its corresponding waiting time is shorter than or equal to the observation time: if $T_i \leq T_{\text{obs}}$, then $X_i = 1$, otherwise $X_i = 0$.

The genotype g_\emptyset is observed if the observation time is shorter than the waiting times of all alterations. Therefore, T_{obs} is the minimum of $n + 1$ competing exponentials, and

$$P(g_\emptyset|\theta_{\text{Null}}) = P(T_{\text{obs}} < \min_{i \in N} T_i) = \frac{\lambda_{\text{obs}}}{\lambda_{\text{obs}} + \sum_{i \in N} \lambda_i} \quad (1)$$

Any genotype g_K with $K \neq \emptyset$ is observed if the waiting times of alterations present in the sample, T_i , for all $i \in K$, are shorter than the observation time, and the waiting times of alterations not present in the sample, T_i , for all $i \in N \setminus K$, are longer than the observation time.

$$P(g_K|\theta_{\text{Null}}) = P(\max_{i \in K} T_i \leq T_{\text{obs}} < \min_{i \in N \setminus K} T_i) \quad (2)$$

The probability that the observation time is shorter than the waiting times of unobserved alterations is not influenced by the specific order between the waiting times of those alterations. Therefore, $P(g_K|\theta_{\text{Null}})$ further equals the sum of the probabilities of all possible specific orders of waiting times of observed alterations. Let $S_K = \{(i_{\sigma(1)}, i_{\sigma(2)}, \dots, i_{\sigma(|K|)}) | i_j \in K \text{ for all } j \text{ and } \sigma \in \Sigma_{|K|}\}$ represent the set of all permutations of indices in K , where Σ_k is the

symmetric group of degree k . By recursively using the expression of the probability of the minimum of competing exponentials, the probability of observing the genotype g_K is

$$\begin{aligned} P(g_K|\theta_{\text{Null}}) &= \sum_{(i_1, i_2, \dots, i_{|K|}) \in S_K} P(T_{i_1} \leq T_{i_2} \leq \dots \leq T_{i_{|K|}} \leq T_{\text{obs}} < \min_{i \in N \setminus K} T_i) \\ &= \sum_{(i_1, i_2, \dots, i_{|K|}) \in S_K} \frac{\lambda_{\text{obs}}}{\lambda_{\text{obs}} + \sum_{i \in N \setminus K} \lambda_i} \prod_{j=1}^{|K|} \frac{\lambda_{i_j}}{\sum_{l=j}^{|K|} \lambda_{i_l} + \lambda_{\text{obs}} + \sum_{i \in N \setminus K} \lambda_i} \end{aligned} \quad (3)$$

As the observations \mathcal{D} contain no temporal information, the model θ_{Null} is unidentifiable.

PROPOSITION 2.1: *The null model $\theta_{\text{Null}} = (\lambda_1, \lambda_2, \dots, \lambda_n, \lambda_{\text{obs}})$ is identifiable only up to λ_{obs} .*

For the proof, see [Supplementary Methods](#). After setting $\lambda_{\text{obs}} = 1$ (without loss of generality), equivalent to scaling the waiting time rates by λ_{obs} , the reparameterized null model $\theta_{\text{Null}} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ becomes identifiable.

2.1.2 Mutual exclusivity model

In the mutual exclusivity model (Fig. 2B), the n genes are assumed to contribute to the same biological function, such that, up to various degrees of mutual exclusivity, only one member is necessary and sufficient to be altered for cancer to progress. An increasing mutual exclusivity interaction in the group directly leads to an increasing fixation probability of a single alteration, corresponding to the gene with the shortest waiting time. The degree of mutual exclusivity of a group of n genes with indices in N , denoted by μ_N , is the probability that the group is perfectly mutually exclusive. μ_N can also be interpreted as the fractional increase in the fixation probability of the genotypes with a single alteration, when more than one gene in the group were altered before observation time, but, due to the mutual exclusivity interaction between the genes in the group, only the one with the shortest waiting time fixates. The fixation of alterations of further genes is suppressed with probability μ_N . Consequently, $1 - \mu_N$ represents the probability of deviating from perfect mutual exclusivity, and for $\mu_N \rightarrow 0$, the mutual exclusivity model is reduced to the null model.

The mutual exclusivity model is parameterized by $\theta_{\text{ME}} = (\lambda_1, \lambda_2, \dots, \lambda_n, \lambda_{\text{obs}}, \mu_N)$. The probability of observing the genotype g_\emptyset is the same as in the null model, as the lack of fixated alterations is uninformative for detecting mutual exclusivity,

$$P(g_\emptyset|\theta_{\text{ME}}) = P(T_{\text{obs}} < \min_{i \in N} T_i) = \frac{\lambda_{\text{obs}}}{\lambda_{\text{obs}} + \sum_{i \in N} \lambda_i} \quad (4)$$

Any genotype g_K with a single alteration, i.e. with $|K| = 1$, can be observed either because K is a mutually exclusive group or because, by chance, the process of tumorigenesis has been observed at the specific point in time when the waiting time of a single alteration was shorter than the observation time. Hence, the probability of observing g_K is the weighted sum of the marginal probability that T_K is the shortest waiting time among all waiting times and the probability that the observed alteration pattern happened in the absence of mutual exclusivity interaction between the genes. The first term represents the probability computed under perfect mutual exclusivity and is weighted by μ_N , while the second represents the

probability computed under the null model and is weighted by $1 - \mu_N$,

$$P(g_K | \theta_{ME}) = \mu_N P(T_K < \min_{i \in N \setminus K} (T_i, T_{obs})) + (1 - \mu_N) P(T_K \leq T_{obs} < \min_{i \in N \setminus K} T_i) \\ = \frac{\lambda_{k; k \in K} \lambda_{obs} + \mu_N \left(\sum_{i \in N} \lambda_i - \lambda_{k; k \in K} \right)}{\lambda_{obs} + \sum_{i \in N} \lambda_i \lambda_{obs} + \sum_{i \in N} \lambda_i - \lambda_{k; k \in K}} \quad (5)$$

Furthermore, observing any genotype g_K with $|K| > 2$, i.e. any genotype with more than one alteration, is considered a deviation from perfect mutual exclusivity. The probability of observing each extra alteration equals the probability that its waiting time is shorter than the observation time, weighted by $1 - \mu_N$, the probability of violating perfect mutual exclusivity,

$$P(g_K | \theta_{ME}) = (1 - \mu_N) P(\max_{i \in K} T_i \leq T_{obs} < \min_{i \in N \setminus K} T_i) \\ = (1 - \mu_N) \sum_{(i_1, i_2, \dots, i_{|K|}) \in S_K} \frac{\lambda_{obs}}{\lambda_{obs} + \sum_{i \in N \setminus K} \lambda_i} \prod_{j=1}^{|K|} \frac{\lambda_{i_j}}{\sum_{l=j}^{|K|} \lambda_{i_l} + \lambda_{obs} + \sum_{i \in N \setminus K} \lambda_i} \quad (6)$$

PROPOSITION 2.2: *The mutual exclusivity model $\theta_{ME} = (\lambda_1, \lambda_2, \dots, \lambda_n, \mu_N, \lambda_{obs})$ is identifiable only up to λ_{obs} .*

For the proof, see [Supplementary Methods](#). Similarly to the null model, after setting $\lambda_{obs} = 1$ (without loss of generality), the reparameterized mutual exclusivity model $\theta_{ME} = (\lambda_1, \lambda_2, \dots, \lambda_n, \mu_N)$ becomes identifiable.

2.1.3 Parameter estimation and testing

The maximum likelihood estimates of all parameters are obtained by setting to zero the corresponding first derivative of the observed log likelihood, numerically approximated using the gradient projection method (Byrd *et al.*, 1995) ([Supplementary Figs S1 and S2](#)). An exception is the case $n = 2$, which allows for an analytical solution for θ_{Null} (the estimates are given in [Supplementary Methods](#)).

PROPOSITION 2.3: *If $n = 2$, then there exists a closed-form solution for the maximum likelihood estimates of λ_1 and λ_2 under the null model θ_{Null} .*

To test for any degree of mutual exclusivity interaction among the n genes, we are testing the alternative hypothesis $\mu_N \neq 0$ versus the null hypothesis $\mu_N = 0$. The logarithm of the ratio of the two likelihoods computed for the maximum likelihood estimates is χ^2 distributed with one degree of freedom (Neyman and Pearson, 1992). The likelihood ratio test statistic is well behaved, as under the null hypothesis, the P values are uniformly distributed ([Supplementary Fig. S3](#)).

2.2 Overall procedure and computational complexity

Our procedure consists of three steps. Given a large dataset of k genes, we first test all $\binom{k}{2}$ pairs for mutual exclusivity, estimating θ_{Null} and θ_{ME} for each pair. The computational complexity of this step is $\mathcal{O}(k^2)$. Second, we construct an undirected graph in which genes are vertices and an edge is drawn between any pair (i, j) if, for chosen thresholds p_{pair} and μ_{pair} , the P value $p_{ij} \leq p_{pair}$ and the degree of mutual exclusivity $\mu_{ij} \geq \mu_{pair}$. The thresholds are chosen based on the sensitivity and specificity levels to which they correspond, as assessed in simulated data. Further, we produce group

candidates by listing all maximal cliques in the constructed graph. To this end, we use the Bron–Kerbosch recursive backtracking algorithm (Bron and Kerbosch, 1973). The upper bound on the running time of the Bron–Kerbosch algorithm is $\mathcal{O}(3^{k/3})$. However, in practice, it is highly efficient (Cazals and Karande, 2008). Finally, we test the candidate groups for mutual exclusivity and select the ones for which the Bonferroni corrected P value is lower than a chosen cutoff. Because of the cubic complexity of matrix inversion (in standard Gauss–Jordan elimination) employed by the numerical optimization routine (Byrd *et al.*, 1995), the complexity of the last step has an upper bound of $\mathcal{O}(s_q q^3)$, where q is the maximal identified clique size and s_q is the number of maximal cliques of size q .

3 Results

3.1 Simulations

We assessed the behavior and performance of TiMEx on simulated data, by varying the waiting time rates of the genes, the degrees of mutual exclusivity of the group and the sample sizes. Specifically, the values of the sample size M were 300, which is similar to the size of the ovarian cancer dataset, 1000, similar to the size of the breast cancer dataset, and 4000, which is a realistic estimate for the size of genomic datasets in the near future. The degrees of mutual exclusivity used for simulations were $\mu \in \{0, 0.2, 0.4, 0.5, 0.6, 0.8, 1\}$. We compared the performance of TiMEx, for both pairs and groups, with a previously introduced permutation-based method by Vandin *et al.* (2012) (ran with 1000 permutations). For the tests on pairs, TiMEx was further compared with one-sided Fisher's exact test for contingency tables, testing whether the number of double mutants is significantly lower than expected under independence. For the tests on larger groups, TiMEx was additionally compared with *muex* (Szczurek and Beerenwinkel, 2014), a previously introduced statistical model for detecting mutually exclusive groups. In a power analysis, we investigated how the sensitivity and specificity of our procedure are influenced by the thresholds on significance and mutual exclusivity degree, p_{pair} and μ_{pair} .

3.1.1 Test performance for pairs and groups

For simulating mutually exclusive gene pairs, we used $\lambda_1, \lambda_2 \in \{0.02, 0.1, 0.5, 1, 3\}$, corresponding to marginal frequencies of the two genes ranging from 2% to 75% in the null model and 0.5% to 75% in the mutual exclusivity model with $\mu_{\{1,2\}} = 1$ ([Supplementary Tables S1 and S2](#)). We performed 100 simulation runs, detected pairwise mutual exclusivity with the three tests and recorded the mean P value ([Fig. 3](#)). When $\mu_{\{1,2\}} = 0$, corresponding to no mutual exclusivity, all three tests do not reject the null hypothesis, with a P value close to 1, for all tested combinations of frequencies and sample sizes. TiMEx is the only test that starts detecting mutual exclusivity from the first non-zero value of $\mu_{\{1,2\}}$ in the chosen simulation set, however with reduced performance for small sample size and small frequencies of both genes. The detection capacity increases with increasing values of $\mu_{\{1,2\}}$, M , λ_1 and λ_2 . For example, for a chosen significance level of 0.05 and a sample size of $M = 4000$, TiMEx detects the gene pairs as being mutually exclusive for any value of $\mu_{\{1,2\}} \geq 0.4$ and for any $\lambda_1, \lambda_2 \geq 0.1$. For higher marginal frequencies, such as corresponding to $\lambda_1, \lambda_2 \geq 0.5$, we can detect mutual exclusivity of degree $\mu_{\{1,2\}} \geq 0.4$ for sample sizes as low as $M = 300$. In contrast, Fisher's exact test and the permutation test in Vandin *et al.* (2012), while performing highly similarly to each other, detect no mutual exclusivity for $\mu_{\{1,2\}} < 0.5$. Moreover, for $\mu_{\{1,2\}} \geq 0.5$, their detection ability is much reduced compared to

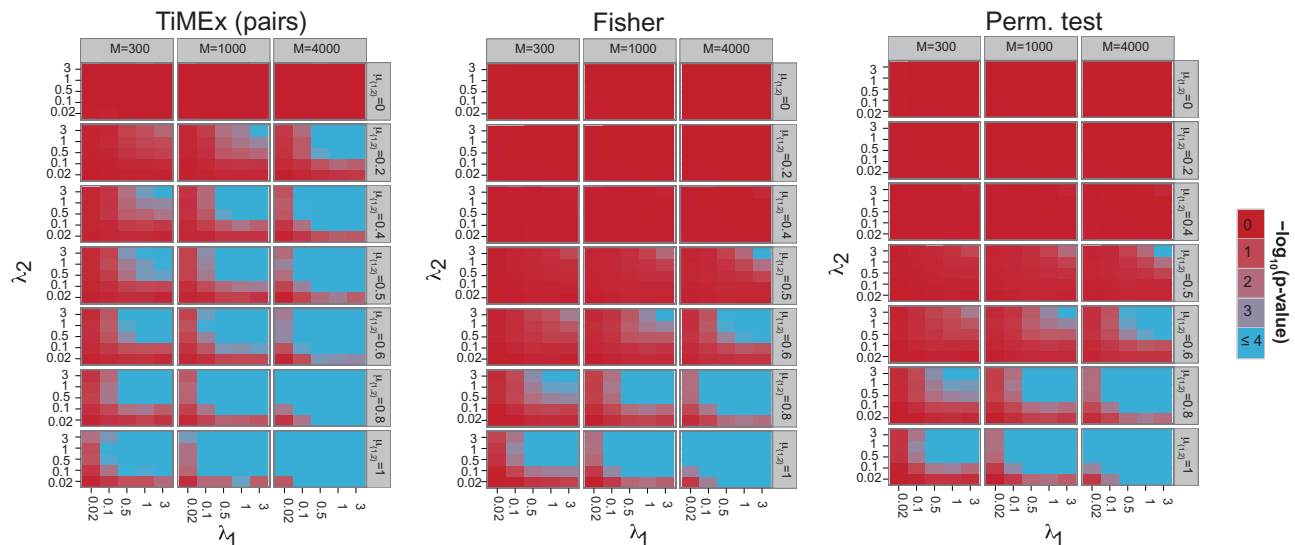


Fig. 3. Mean P value (over 100 simulation runs) for TiMex, Fisher's exact test and the permutation test in Vandin et al. (2012), for different sample sizes M , pairwise degrees of mutual exclusivity $\mu_{1,2}$ and waiting time rates determining the marginal frequencies of the two genes, λ_1 and λ_2 . TiMex is highly sensitive in detecting mutual exclusivity and outperforms the other two tests, which only start detecting mutual exclusivity for $\mu_{1,2} \geq 0.5$. The detection capacity increases with increasing values of $\mu_{1,2}$, M , λ_1 and λ_2

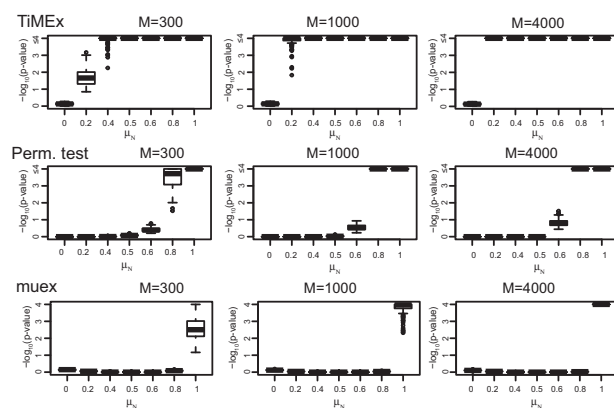


Fig. 4. Summary P value (over 100 simulation runs and 10 simulated groups of size 5) of TiMex, the permutation test in Vandin et al. (2012) and muex (Szczurek and Beerenwinkel, 2014), for different sample sizes M and degrees of mutual exclusivity μ_N . TiMex is highly sensitive in detecting mutual exclusivity and outperforms both other methods. The permutation test only detects mutual exclusivity for $\mu_N \geq 0.8$ and outperforms muex, which only detects pure mutual exclusivity ($\mu_N = 1$). The detection ability of TiMex improves with increasing values of M and μ_N

TiMex. The null model used in Fisher's exact test is a classical independence model, while the waiting times in our null model are not statistically independent, even with fixed rate of the observation time λ_{obs} .

For simulating mutually exclusive groups, we fixed the group size to 5 and produced 10 different groups by uniformly sampling waiting time rates with values between 0.01 and 1, which correspond to an expected alteration frequency of 14% (Supplementary Table S3). We performed 100 simulation runs, detected mutual exclusivity with TiMex, the permutation test and muex and summarized the P value over both simulated groups and simulation runs (Fig. 4). Similar to the case of pairs, the detection ability of TiMex increases with increasing sample size and degree of mutual exclusivity. For a significance level of 0.05, we detect mutual exclusivity for

almost all tested sample sizes and degrees of mutual exclusivity, with the exception of small sample size $M = 300$ and low degree of mutual exclusivity $\mu_N = 0.2$. For the highest sample size, $M = 4000$, TiMex is very sensitive in detecting mutual exclusivity for any tested positive degree, with mean P value $\leq 10^{-4}$. In contrast, the permutation test only starts detecting mutual exclusivity for $\mu_N \geq 0.8$, however, outperforming muex, which only detects pure mutual exclusivity ($\mu_N = 1$). On data simulated using $\mu_N = 0$, all three tests do not reject the null hypothesis with mean P value > 0.6 .

In addition to assessing the detection ability of TiMex and other methods on data simulated from TiMex, we conducted simulations on datasets generated in a more general manner. We generated groups of $n = 5$ mutually exclusive genes by varying the sample size as before, the coverage, i.e. the percentage of tumors with one alteration, among $\{0.2, 0.4, 0.6, 0.8\}$, and the probability of passenger alterations among $\{0.001, 0.01, 0.1\}$. On all datasets, TiMex outperforms the permutation test and muex and always records lower P values (ranking not shown for P values lower than 10^{-10}) (Supplementary Fig. S4). All three methods perform better with increasing sample size, increasing coverage and decreasing passenger probability, and, for most of the tested values, they significantly detect the group as mutually exclusive. For TiMex, we also estimated the degree of mutual exclusivity μ_N corresponding to the generated groups (Supplementary Fig. S5). For the very low passenger probability of 0.001, the inferred degree of mutual exclusivity is 1, as the expected number of passenger mutations per dataset is very low, especially for small sample sizes. The lowest inferred degree of mutual exclusivity is 0.7, corresponding to small coverage and small sample size. The estimated μ_N increases with increasing coverage and decreasing passenger probability, and the estimation improves with increasing sample size.

3.1.2 Power analysis

For assessing the true- and false-positive rates of our procedure, we constructed 100 datasets consisting of two groups: a group of size 3 simulated from the mutually exclusive model and a group of size 9 simulated from the conditionally independent model. The

corresponding λ values were sampled uniformly between 0.01 and 1 (Supplementary Table S3). We tested all pairs with TiMEx, detected maximal cliques as candidates and evaluated them with TiMEx. An identified group was considered to be mutually exclusive if its Bonferroni corrected P value was lower than 0.1. We computed the true-positive rate by counting a single time, among the mutually exclusive groups of size at least 3, all edges only connecting two genes part of the true mutually exclusive group and normalizing by the number of all possible such edges. Similarly, we computed the false-positive rate by counting a single time, among the mutually exclusive groups of size at least 3, all edges not connecting two genes part of the true mutually exclusive group and normalizing accordingly. TiMEx performs generally very well in reconstructing the implanted mutually exclusive group (Supplementary Fig. S7). The highest effect in increasing the true-positive rate and decreasing the false-negative rate is given by increasing the mutual exclusivity degree of the simulated group. For values of the threshold $\mu_{\text{pair}} \geq 0.5$, the false-positive rate is often set to 0, and most of the times reduced by at least 75% when compared with the case when $\mu_{\text{pair}} = 0.2$. However, for small degrees of mutual exclusivity of the simulated group, both the true and the false-positive rates are often simultaneously reduced.

Additionally, on the same simulated datasets, we analyzed how often the true mutually exclusive group is ranked first by corrected P value among the identified groups (Supplementary Fig. S8). The threshold μ_{pair} largely impacts the detection performance, while the impact of p_{pair} is negligible. For high degrees of mutual exclusivity of the true group, $\mu_N \geq 0.8$, the real group is either top ranked or a strict subset of the top ranked one, depending on μ_{pair} . Optimal performance is achieved for $\mu_{\text{pair}} = 0.5$, corresponding to a percentage of between 50% and 100% of datasets for which the true group is top ranked. Moreover, if the true group is perfectly mutually exclusive ($\mu_N = 1$), it is top ranked in more than 90% of the datasets for medium values of μ_{pair} , for any value of p_{pair} and for medium sample sizes. For lower values of μ_{pair} , the true group is a strict subset of the top ranked group. For low degrees of mutual exclusivity μ_N , either no significant groups are identified or the top ranked group is not the real group. The detection power improves with increasing sample size and increasing degree of mutual exclusivity μ_N .

3.2 Biological datasets

We ran our procedure on four biological datasets: two glioblastoma datasets preprocessed by muex (Szcurek and Beerenwinkel, 2014) and Multidendrix (Leiserson et al., 2013) and two datasets downloaded from TCGA and preprocessed as explained in Supplementary Section S2.3: breast cancer and ovarian cancer. Our main interest was detecting gene groups with average or high degree of mutual exclusivity and minimizing the false-positive rate, while maintaining the true-positive rate at a high level. Therefore, based on the sensitivity and specificity estimates in simulated data (Supplementary Figs S7 and S8), we set $\mu_{\text{pair}} = 0.5$ and $p_{\text{pair}} = 0.01$ for the four datasets. A detected group was considered significantly mutually exclusive if its Bonferroni-corrected P value (q value) was less than 0.1 (Supplementary Fig. S9). To test the stability of the identified groups, we subsampled the set of patients at frequencies of 30%, 50% and 80% and repeated the procedure 100 times, reporting how often each group is still identified as mutually exclusive (Supplementary Tables S4–S15). Among the identified groups of any size, we further computed the most stable subgroups (Supplementary Tables S16–S27). For mutually exclusive groups with high enough alteration frequencies, higher stability indicates

stronger mutual exclusivity support in the data. For each group size, we tested the first 10 groups ranked by q value for pathway enrichment with WebGestalt (Zhang et al., 2005) on the Pathways Commons dataset (Cerami et al., 2011) and reported all significantly enriched pathways for a BH-corrected P value threshold of 0.01. On all four datasets, we compared our results with two other methods: Multidendrix and muex (Supplementary Tables S32–S37). Supplementary Section S1 discusses the mutually exclusive groups identified in the two glioblastoma datasets.

3.2.1 Mutual exclusivity in breast cancer

In the breast cancer dataset (Fig. 5), we found 63 groups of size 2, 416 groups of size 3, 96 groups of size 4 and 10 groups of size 5. Since all the 10 largest groups contained one gene with frequency less than 10%, these groups were highly unstable to subsampling, even if they corresponded to functionally related collections of genes (Supplementary Table S13). The first three groups with the lowest q value consisted of the point mutations of the tumor suppressors *CDH1*, *GATA3*, *MAP3K1*, the copy number aberration of *CDKN1B*, which belong to pathways including PI(3)K, mTOR, PDGF receptor signaling network or EGF receptor (ErbB1) signaling and the copy number aberration of one of the following genes: *MIEN1*, *PPP1R1B* or *ERBB2*. *MIEN1* is an oncogenic protein, whose overexpression functionally enhances migration and invasion of tumor cells via modulating the activity of the PI(3)K pathway (Hsu et al., 2012), providing evidence for the functional relation between these genes. Moreover, the *PPP1R1B*-*STARD3* chimeric fusion transcript was shown to activate the PI(3)K/AKT signaling pathway and promote tumorigenesis (Yun et al., 2014), while *ERBB2* is an oncogene that also belongs to the PI(3)K and mTOR pathways. The next two mutually exclusive groups of size 5 included the same three point mutations, and the copy number aberrations of either *MIEN1* and *B4GALNT3*, which have no known functional

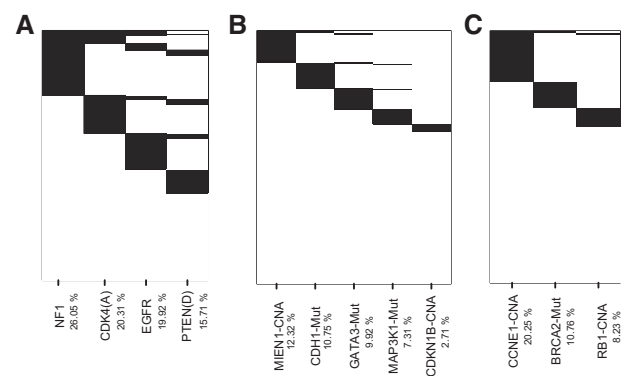


Fig. 5. The alteration frequencies of selected identified mutually exclusive groups. The horizontal axis displays the members of each group, together with their relative frequency in the dataset, as well as their alteration type (Mut for mutation and CNA for copy number aberration). A black line is drawn whenever an alteration is present in a sample. (A) The group consisting of the deletion of *PTEN*, the amplification of *CDK4* and the point mutations *EGFR* and *NF1* (q value 6e-10) was the most stable among the groups of largest size identified by TiMEx in the glioblastoma dataset used by Multidendrix in Leiserson et al. (2013). (B) The group consisting of the point mutations of *CDH1*, *MAP3K1*, *GATA3* and the copy number aberrations of *CDKN1B*, *MIEN1* (q value 1e-23) was the most significant group of largest size identified by TiMEx in the breast cancer dataset. (C) The group consisting of the point mutation of *BRCA2* and the copy number aberrations of *RB1* and *CCNE1* (q value 5e-09) was the most significant and most stable group of largest size identified by TiMEx in the ovarian cancer dataset

role in breast cancer, or *GRB7* and *CDKN1B*, which are part of the *Common group of pathways*. The group of size 4 with lowest q value consisted of the point mutations of *CDH1*, *MAP3K1*, *TP53* and *GATA3* and was entirely mapped to the *Common group of pathways*, as well as to the CDC42 signaling events pathway (Supplementary Table S12). The second and the third group included, instead of the *GATA3* point mutation, the copy number aberration of either *TUBD1* or *INTS4*. Even though strong evidence of association for these two genes and the group of three point mutations exists in the data, *TUBD1* and *INTS4* have no known functional role in cancer. Furthermore, no known functional connections were found between the identified pairwise mutually exclusive interactions (Supplementary Table S10), mainly because most of the functionally related pairs were part of cliques of larger sizes. The subgroups with highest subsampling stability (Supplementary Tables S22–S24) consisted of genes with known functional involvement in cancer and part of cancer pathways, such as *GATA3*, *PIK3CA*, *PTEN* or *MIEN1*.

We separately ran our procedure on the subset consisting of 507 samples annotated as distinct breast cancer subtypes (Supplementary Tables S28–S31). None of the alterations identified in the top ranking groups were specific to the Basal subtype (Supplementary Table S28), while some of the top ranked mutually exclusive relations identified based on the entire dataset were identified in subsets of data for Her2, LuminalA and LuminalB subtypes. For example, the mutually exclusive connections between the point mutations of *PIK3CA* and *GATA3*, as well as between the point mutations of *MAP3K1*, *GATA3* and *TP53* were identified in the LuminalA subtype (Supplementary Table S30). Similarly, the group including the point mutations *PIK3CA*, *TP53* and *GATA3* was specific to LuminalB (Supplementary Table S31). The point mutations of *PIK3CA* and *TP53*, and the copy number aberration of *PTEN* were part of multiple mutually exclusive pairs identified in the Her2 subtype (Supplementary Table S29). *TUBD1*, a gene part of multiple groups, was mutually exclusive with the point mutation of *MAP3K1* in LuminalA and with the point mutations of *PIK3CA* and *TP53* in LuminalB.

We ran Multidendrix using $\alpha = 2.5$ [as suggested in Leiserson et al. (2013)], $t = 4$ and a range of k_{\max} values (Supplementary Table S35). Multidendrix primarily identified the core group including the point mutations of *TP53*, *GATA3* and *MAP3K1*, however, in the same group as the point mutations of *CTCF* and *PLXNB2*, which do not belong to any known functional pathways. On the contrary, TiMex identified these three point mutations in a common module with the point mutation of *CDH1*. Similarly, the next two modules with highest weight only contained three genes in known functional pathways, as assessed by WebGestalt, on the Pathway Commons database (data not shown). The fourth module identified by Multidendrix contained no significant pathways. *muex* did not scale to the size of this dataset, and none of the top 30 groups of any size identified by TiMex were found significantly mutually exclusive by *muex*'s statistical test.

3.2.2 Mutual exclusivity in ovarian cancer

In the ovarian cancer dataset (Fig. 5), we identified 25 mutually exclusive groups of size 2 and 19 groups of size 3. The top ranked group of size 3 (Supplementary Table S15) included genes part of the FOXM1 transcription factor network, and involved in cell cycle regulation, recently shown to play a major role in the progression of ovarian cancer (Cancer Genome Atlas Research Network, 2011): the copy number aberrations of the tumor suppressor gene *RB1* and

the oncogene *CCNE1* and the point mutation of the tumor suppressor gene *BRCA2*. This group was also highly stable to subsampling (identified in 92% of the cases in which 80% of the patients were subsampled). The other top ranked group whose members mapped entirely to known pathways consisted of the point mutations of *BRCA1* and *BRCA2*, and the copy number aberration of *CCNE1*, which are core members of the ATM pathway. These two modules have also been previously identified as mutually exclusive by MEMo (Ciriello et al., 2012). The second and third top scoring groups of size 3 included the copy number aberrations of *MYC* and *CCNE1*, two members of cell cycle regulation pathways involved in the G1/S phase transition, also identified by MEMo, together with the copy number aberration of one gene with yet unknown functional role in ovarian cancer: *WNK1* or *NINJ2*. The top ranked mutually exclusive pair, which was also highly stable to subsampling (identified 98% of the times when subsampling 80% of the patients) included *KRAS* and *TP53* point mutations, which are part of the p75 NTR receptor-mediated signaling pathway (Supplementary Table S14). The second mutually exclusive pair, also highly stable to subsampling, included the point mutations of *TP53* and *RB1*, both part of the TGFBR and p53 pathways. Further mutually exclusive connections stable to subsampling included the point mutations of *BRCA1* and *BRCA2* and the copy number aberration of *CCNE1* (Supplementary Tables S25–S27).

We ran Multidendrix using $t = 4$ and a range of k_{\max} values (Supplementary Table S36). The groups identified by TiMex and Multidendrix showed a high overlap. For example, the top ranking groups identified by TiMex, i.e. the pair including the point mutations of *TP53* and *KRAS*, and the group including the copy number aberrations of *RB1* and *CCNE1*, together with the point mutation of *BRCA2*, were also identified by Multidendrix. Moreover, subsets of most of the group members that Multidendrix identified, for e.g. $k_{\max} = 5$ were identified by TiMex as groups of size 3, such as the point mutations of *BRCA1* and *BRCA2* and the copy number aberration of *EPHX3*. Even though *muex* did not scale to exhaustively analyze the dataset, 14 of the pairs and 4 of the groups of size 3 identified by TiMex were found to be significant by *muex* (Supplementary Table S37). Almost all the pairs included either the point mutation of *BRCA2* or the copy number aberration of *NF1*, while the larger groups included genes mapping to relevant pathways, among which many had also been identified by Multidendrix. The reason why these groups are also found to be mutually exclusive by *muex* is the fact that the alteration frequencies of their members are balanced.

4 Discussion

We have introduced TiMex, a probabilistic generative model for detecting mutual exclusivity patterns of various degrees across carcinogenic alterations, and an efficient multistep procedure for identifying all mutually exclusive groups in large datasets. TiMex is the first method that describes the mutual exclusivity property as a consequence of a dynamic process in time. Unlike previous *de novo* approaches, TiMex infers functional relations between genes based on an underlying temporal representation of the process of gene alteration in tumorigenesis. Moreover, TiMex is a probabilistic generative model, providing a natural way of rigorously quantifying the degree and significance of mutual exclusivity of a group of genes. To the best of our knowledge, TiMex is the first method inferring a continuous range of mutual exclusivity degrees. Biologically, the small, but observable, increase in tumor fitness due to multiple

alterations in a group of functionally related genes supports the hypothesis that mutual exclusivity occurs at various degrees, as opposed to a binary classification. Unlike most other approaches, TiMEx does not explicitly impose frequency constraints and detects both high frequent and very low frequent alterations, solely based on the temporal relations between them. It identifies all mutually exclusive gene groups of various, not pre-defined sizes and performs highly efficiently on large datasets. In simulation studies, TiMEx outperforms previous methods for detecting mutual exclusive groups, showing high sensitivity even at low degrees of mutual exclusivity and scaling very well to sample sizes of several thousands tumors, which is expected to be soon reached by cancer genome sequencing studies. On biological datasets, most of the top ranked mutually exclusive groups identified by TiMEx have stronger functional biological relevance than the groups identified by previous methods. These results clearly indicate that TiMEx is not only theoretically justified by its biological and probabilistic foundation in describing tumorigenesis as a generative process of mutually exclusive alteration patterns, but is also efficiently and fruitfully applicable in practice.

TiMEx models, however, still a simplified representation of carcinogenesis. Given a particular order between the waiting times of the genes and the observation time, the probability of violating mutual exclusivity, $1 - \mu_N$, is independent of how many or which alterations are in a group. One natural extension of TiMEx would be to consider an incremental penalty for additional point alterations violating perfect mutual exclusivity, hence increasing the probability of being in a nonmutually exclusive state with increasing number of violating alterations. Additionally, even if highly efficient, the search for mutually exclusive gene groups is heuristic and depends on the thresholds p_{pair} and μ_{pair} . With overly stringent thresholds, too few candidates would be proposed, while using overly permissive thresholds would lead to selecting as candidates a vast number of subsets, making the procedure intractable. Moreover, the fact that pairwise connections are selectively included when testing groups for mutual exclusivity, on the basis of the same data, renders the interpretation of the resulting P values more complex. To address this, we propose setting the thresholds following the desired sensitivity-specificity tradeoff as assessed in detail in the simulation studies discussed here. Finally, ranking the identified mutually exclusive groups by the resulting q values facilitates the interpretation of their functional role in carcinogenesis.

The exponential distribution, used for modeling the waiting times to alterations and the observation time, is a typical choice to describe waiting times (Gerstung and Beerenwinkel, 2010), both due to its generality and to its mathematical convenience. While the exponential distribution is the simplest model for system failure time, other families of distributions can be readily integrated into our mathematical framework, with nevertheless the cost of more involved mathematical formulas. For example, using the Weibull distribution provides a supporting assumption in modeling cancer progression due to the fact that the instantaneous probability of occurrence of an event changes with time. Another extension of TiMEx is renouncing to the independence assumption at the level of observations and applying our procedure to large-scale time series data of tumor progression. Once this type of data becomes available,

TiMEx will facilitate a more detailed understanding of pathways involved in tumor progression.

Acknowledgement

The authors would like to thank Hesam Montazeri, David Seifert and Jack Kuipers for useful discussions and suggestions.

Funding

Simona Constantinescu was financially supported by the Swiss National Science Foundation (Sinergia project 136247).

Conflict of Interest: none declared.

References

- Bron, C. and Kerbosch, J. (1973) Algorithm 457: finding all cliques of an undirected graph. *Commun. ACM*, **16**, 575–577.
- Byrd, R.H. *et al.* (1995) A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.*, **16**, 1190–1208.
- Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature*, **474**, 609–615.
- Cazals, F. and Karande, C. (2008) A note on the problem of reporting maximal cliques. *Theor. Comput. Sci.*, **407**, 564–568.
- Cerami, E.G. *et al.* (2011) Pathway commons, a web resource for biological pathway data. *Nucleic Acids Res.*, **39**(Suppl. 1), D685–D690.
- Ciriello, G. *et al.* (2012) Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res.*, **22**, 398–406.
- Ding, L. *et al.* (2008) Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*, **455**, 1069–1075.
- Gerstung, M. and Beerenwinkel, N. (2010) Waiting time models of cancer progression. *Math. Popul. Stud.*, **17**, 115–135.
- Hsu, C.-H. *et al.* (2012) Solution structure of the oncogenic MIEN1 protein reveals a thioredoxin-like fold with a redox-active motif. *PLoS One*, **7**, e52292.
- Jerby-Arnon, L. *et al.* (2014) Predicting cancer-specific vulnerability via data-driven detection of synthetic lethality. *Cell*, **158**, 1199–1209.
- Leiserson, M.D. *et al.* (2013) Simultaneous identification of multiple driver pathways in cancer. *PLoS Comput. Biol.*, **9**, e1003054.
- Miller, C.A. *et al.* (2011) Discovering functional modules by identifying recurrent and mutually exclusive mutational patterns in tumors. *BMC Med. Genomics*, **4**, 34.
- Neyman, J. and Pearson, E.S. (1992) *On the Problem of the Most Efficient Tests of Statistical Hypotheses*. Springer, New York.
- Sjöblom, T. *et al.* (2006) The consensus coding sequences of human breast and colorectal cancers. *Science*, **314**, 268–274.
- Szczurek, E. and Beerenwinkel, N. (2014) Modeling mutual exclusivity of cancer mutations. *PLoS Comput. Biol.*, **10**, e1003503.
- Vandin, F. *et al.* (2012) De novo discovery of mutated driver pathways in cancer. *Genome Res.*, **22**, 375–385.
- Yeang, C.-H. *et al.* (2008) Combinatorial patterns of somatic gene mutations in cancer. *FASEB J.*, **22**, 2605–2622.
- Yun, S. *et al.* (2014) PPP1R1B-STARD3 chimeric fusion transcript in human gastric cancer promotes tumorigenesis through activation of PI3K/AKT signaling. *Oncogene*, **33**, 5341–5347.
- Zhang, B. *et al.* (2005) WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.*, **33**(Suppl. 2), W741–W748.