FOOD COMPOSITION AND ADDITIVES

GMOtrack: Generator of Cost-Effective GMO Testing Strategies

PETRA KRALJ NOVAK
Jožef Stefan Institute, Department of Knowledge Technologies, Jamova 39, 1000 Ljubljana, Slovenia
KRISTINA GRUDEN and DANY MORISSET
National Institute of Biology, Department of Biotechnology and Systems Biology, Večna pot 111, 1000 Ljubljana, Slovenia
NADA LAVRAČ
Jožef Stefan Institute, Department of Knowledge Technologies, Jamova 39, 1000 Ljubljana, Slovenia and University of Nova Gorica, Vipavska 13, 5000 Nova Gorica, Slovenia
DEJAN ŠTEBIH, ANA ROTTER, and JANA ŽEL
National Institute of Biology, Department of Biotechnology and Systems Biology, Večna pot 111, 1000 Ljubljana, Slovenia

Commercialization of numerous genetically modified organisms (GMOs) has already been approved worldwide, and several additional GMOs are in the approval process. Many countries have adopted legislation to deal with GMO-related issues such as food safety, environmental concerns, and consumers' right of choice, making GMO traceability a necessity. The growing extent of GMO testing makes it important to study optimal GMO detection and identification strategies. This paper formally defines the problem of routine laboratory-level GMO tracking as a cost optimization problem, thus proposing a shift from "the same strategy for all samples" to "sample-centered GMO testing strategies." An algorithm (*GMOtrack*) for finding optimal two-phase (screening–identification) testing strategies is proposed. The advantages of cost optimization with increasing GMO presence on the market are demonstrated, showing that optimization approaches to analytic GMO traceability can result in major cost reductions. The optimal testing strategies are laboratory-dependent, as the costs depend on prior probabilities of local GMO presence, which are exemplified on food and feed samples. The proposed *GMOtrack* approach, publicly available under the terms of the General Public License, can be extended to other domains where complex testing is involved, such as safety and quality assurance in the food supply chain.

In 2008, after more than a decade of commercialization of genetically modified (GM) crops, the total accumulated land areas sown with GM plants have exceeded 800 million hectares. Between 1996 and 2008, a total of 670 approvals in 55 countries were granted for 144 events for 24 crop species (1). Not only the presence of genetically modified organisms (GMOs) on the market, but also the diversity of their taxonomy (taxon host plants) and biotechnology (genetic constructs) is increasing. More than 40 countries worldwide have adopted legislation dealing with GM crop-related issues. The goal of these regulations is primarily to ensure food safety. These policies also ensure the consumers' right of choice between GMO and non-GMO-derived products (2). The coexistence of conventional, organic, and GM crop production also requires additional traceability and fulfillment of legal obligations concerning possible economic implications of GMO admixture (2, 3). Accurate detection and identification of GMOs are therefore necessary to ensure control of GMO traceability. In most GMO detection enforcement laboratories, the standard testing technology is PCR or its quantitative derivative, real-time PCR. Different DNA elements can be targeted by PCR. On the one hand, screening elements that are found in several different GMOs can be used to detect the presence of GMOs, and on the other hand, event-specific targets are used for the identification of individual GMOs (4). The reference gene system targets can also be used to determine which host plants are present in the sample.

Currently used GMO testing strategies in routine GMO testing laboratories consist of two phases: a screening phase followed by an identification phase (Figure 1). In the screening phase, GMO presence is detected by performing a minimal number of screening assays that cover all GMOs in question. If GMO presence is indicated, a second phase using event-specific assays is performed to identify possible GMOs in the sample. In the screening phase, a minimal number of most common genetic elements found in GM crops are used (4), thereby minimizing the screening cost and not the overall analysis cost. The increasing presence of GMOs on the market and their growing taxonomic and biotechnological diversity are rendering the second phase of current GMO testing strategies unmanageable both in terms of time and associated testing costs. Consequently, it is necessary either to

Received February 18, 2009. Accepted by SG May 13, 2009.
Corresponding author's e-mail: petra.kralj.novak@ijs.si

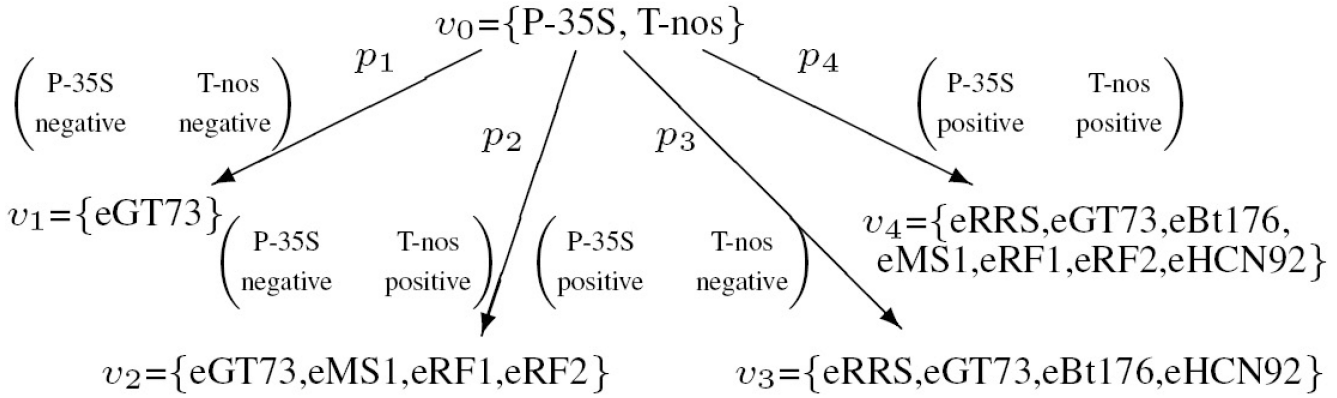

Figure 1. An example testing strategy. The first testing phase includes P-35S assay and T-nos assay. The second phase, depending on the outcome of the first phase, has between one and seven event-specific assays.

introduce new analytical technologies for high-throughput GMO diagnostics or to develop new tools to support decision making in determining cost-effective GMO testing strategies, including optimized GMO identification.

In this paper, we present a new sample-centered approach to analytical GMO traceability. We propose that, instead of using the same testing strategy regardless of the characteristics of the sample, the testing strategy should be tuned to the given sample being analyzed in order to minimize the total analysis cost. To study the benefits of the sample-centered GMO testing approach, we have (*1*) gathered relevant information on GM crops in a database that combines data from various sources and includes innovative information on the frequencies of appearance of individual GMOs; (*2*) developed and implemented an algorithm called *GMOtrack*, which generates testing strategies that use a combination of screening assays that best trade off the screening and the expected event-specific costs; and (*3*) experimentally evaluated this approach for generating cost-effective GMO testing strategies.

Methods

GMOtrack: The Formal Background

GMOs can be described with combinations of elements (promoters, terminators, genes, and other genetic elements) inserted into the host plant genome. Therefore, it is possible to use methods specifically targeting these elements in routine detection to design a smart screening phase before performing the final identification and quantification of the GMOs. The *GMOtrack* approach uses the following assumptions:

The assays' outcome is either positive or negative, and their accuracy is satisfactory (there are no false-positive or ambiguous results); the negative screening assay results confirm the absence of the GMOs which they are able to detect; only the use of an appropriate event-specific assay allows confirmation of the presence of a specific GMO; we assume that a complete set of event-specific assays is at our disposal.

The technology used for testing is real-time PCR and the costs of analysis are formulated accordingly. The cost of all assays is the same, but the testing cost per assay decreases with an increasing number of assays applied simultaneously. GMOs are independent. Knowing that one GMO is present in the sample provides no information about the presence of other GMOs in the same sample.

Strategies generated by *GMOtrack* are similar yet different than current testing strategies. On the one hand, both types of strategies have two phases: a screening phase and an identification phase. In the identification phase, the event-specific assays for all possible GMOs according to the result of the screening phase are performed. On the other hand, unlike current testing strategies which minimize the number of screening assays, the *GMOtrack* testing strategies propose a combination of screening assays that best trade off the screening and the expected event-specific costs.

Results

The benefits of sample-centered GMO testing strategies are shown in two use cases. First, the need for cost optimization with increasing GMO presence on the market is demonstrated. Second, optimal testing strategies are shown to differ according to prior probabilities of GMO presence, exemplified on food and feed samples.

Data Acquisition

Data from several publicly available databases were collected to build the working database on GM crops and their genetic elements (5–10). Only data about GM crops relevant to the food and feed legislation were included in the database. We have gathered the frequencies of presence of each GMO per sample in different years, according to the actual situation in our routine GMO testing laboratory (termed "TestLab" in the following). If a GMO was never detected, it was assigned a low frequency of one per thousand (0.001).

All the data were preprocessed to avoid false-positive results. The data were also preprocessed to remove redundancy. First, screening elements that detect one GMO were removed, because in testing strategies the suitable event-specific elements should always be used instead. Second, if two or more screening assays detect the same set of

Table 1. Sample dataset of preprocessed GMO data[a]

			Screening assays (T_S)								Event-specific assays (T_E)					
GMO	Crops	Probability	P-35S	P-TA29	P-nos	CP4 epsps	bar	Barstar	T-nos	T-35S	P-35S::bar	RRS	GT73	Bt176	MS1	RF1	RF2	HCN92
RRS	Soybean	0.03	x			x			x			x						
GT73 (RT73)	Oilseed rape	0.001				x							x					
Bt176	Maize	0.001	x				x			x	x			x				
MS1	Oilseed rape	0.001		x	x		x		x		x				x			
RF1	Oilseed rape	0.001		x	x		x	x	x		x					x		
RF2	Oilseed rape	0.001		x	x		x	x	x								x	
HCN92	Oilseed rape	0.001	x		x					x								x

[a] A sample dataset of preprocessed GMO data describing authorized GMOs in the European Union in the year 1997. Each line is one GMO; columns are screening assays targeting genetic elements inserted in the GM crop followed by event-specific assays allowing precise identification of a unique GMO. The presence of genetic elements is indicated by marks in the table, meaning that the organism in the corresponding row should react positively to an assay specific to the genetic element in the corresponding column. Probability denotes the frequency of GMO appearance in the TestLab in the year 1997.

GMOs (identical columns in the table), they form equivalence classes. One representative of each equivalence class remained in the table because, in such a case, one assay provides the same information as the other(s) in the equivalence class.

A sample dataset, describing GMOs approved in 1997, is shown in Table 1. All the collected data, assembled by years, are available at http://kt.ijs.si/software/GMOtrack/. Additionally, and to evaluate the usability of our framework in the context of increasing number and variety of GMOs, we have simulated a future situation (called "beyond 2008") where a sample should be analyzed over a set of GMOs that are the GM crops being approved for use in food and feed in the European Union (EU) in April 2008, plus additional GM crops with applications awaiting approval.

Experimental Evaluation

The applicability of the proposed laboratory-level GMO traceability cost optimization is shown in two use cases. In the first use case, we investigated optimal testing strategies in time, from 1996, when the first GMO was introduced on the EU market, to the simulated situation beyond 2008, when identification of 37 GMOs might be necessary. In the second use case, we investigated how optimal testing strategies vary if sample-dependent probabilities of GMO presence are known.

We used the *GMOtrack* algorithm to compute the optimal two-phase testing strategies. We approximated the *g* function (cost of one run of PCR assays) for the chosen laboratory with a linear function that depends on the number of assays (*numAssays*) according to the equation $g = 21.18 \cdot numAssays + 91.82$. It is a simplification of the real situation with a relative absolute error of 3%.

GMO Traceability Costs in Time

The aim of this use case is to show how GMO identification costs per sample increase in time with increased diversity and abundance of GMOs on the market. The effect is much less pronounced if optimal testing strategies are used.

We collected data about GMOs authorized in the EU from 1996, when the first GMO [Roundup® Ready soybean, (RRS)] was authorized on the EU market, up to 2008, when 22 GMOs were authorized, along with their probabilities. To estimate the GMO identification costs in a situation simulated beyond 2008, we added to the data from 2008 additional data on 15 GMOs that are in the pipeline for approval in the EU. A conservative approximation of probabilities of GMOs appearing in samples beyond 2008 was used. The probabilities of GMOs authorized by 2008 stay the same, while the probabilities of the other (new) GMOs are set to $1^0/_{00}$.

We compared the strategies generated by *GMOtrack* with the strategy actually used in the TestLab, which is a detection strategy with a single screening step combining the P-35S and T-nos screening assays and the RT73 event-specific assay. The results of our experiment are shown as trend lines in Figure 2. The underlying data are presented in Table 2. The total expected cost is the sum of the screening and the

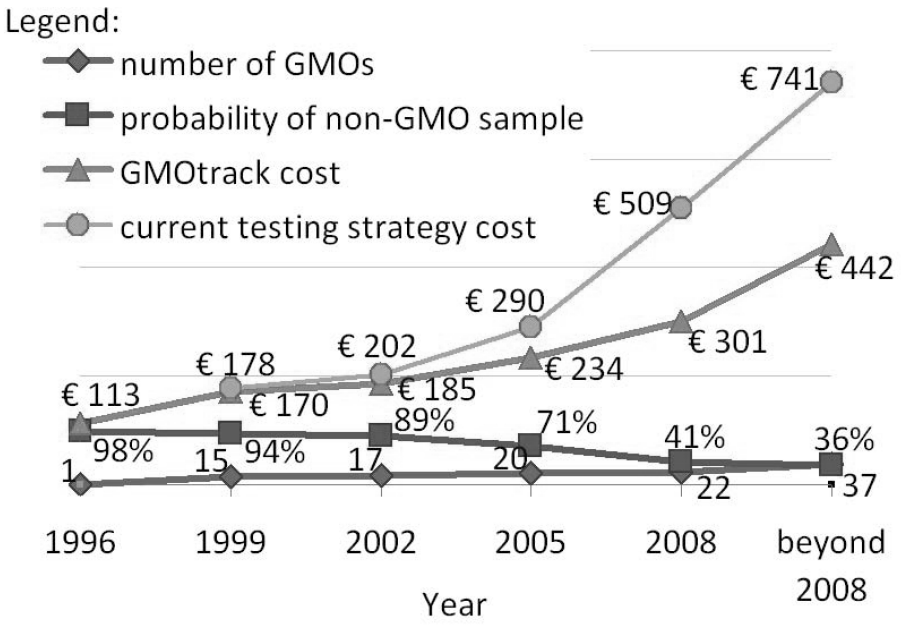

Figure 2. Comparison of the current and *GMOtrack* GMO testing strategies. Evolution of GMO testing costs from 1996 to 2008 and beyond, presenting the trend lines for the number of authorized GMOs, probability of non-GMO sample and expected costs for the optimal two-phase strategy compared to the current testing strategy cost. Costs were calculated for the real-time PCR technology, based on the TestLab data. Data for "beyond 2008" were simulated by adding the GMOs with awaiting authorizations to the GMOs that were authorized by April 2008.

expected event-specific costs for the optimal number and combination of screening assays. Trend lines representing the number of authorized GMOs in the EU and the prior probability of a sample being non-GMO are also displayed.

Optimal Testing Strategies for Different Sample Types

We investigated the optimal testing strategies if sample-dependent probabilities of GMO presence are known. For this purpose, probabilities for 2008 were estimated separately for two types of samples: food and feed (as opposed to single probabilities in the data in Table 1).

The probability estimates of GMOs in food samples are as follows: the probability of RRS soybean is 14%, and the probability of 21 other GMOs is $1^0/_{00}$. On the other hand, the probability estimates of GMOs in feed samples are as follows: probability of RRS soybean is 54%, probability of maize (GA21, Bt176, Bt11, NK603, and MON810) is 2% each, and the probability of 16 other GMOs is $1^0/_{00}$. From these data, we can directly compute the probability of a sample being non-GMO. The prior probability of a food sample being non-GMO is 82%, and the prior probability of a feed sample being non-GMO is 40%.

The *GMOtrack* testing strategies for the two sample types, including their costs, are shown in Table 3, where the expected total cost is the sum of the screening and identification costs (rounded to increase readability). Screening costs depend only on the number of screening assays, whereas the expected identification cost is computed according to the probabilities of screening outcomes and the number of event-specific assays needed in the second phase. Setting the probabilities of GMO presence according to the knowledge about the sample allows *GMOtrack* to tune the testing strategy to the sample.

From the results presented in Table 3, we can draw the following conclusions. First, optimal GMO testing strategies depend on probabilities of GMO presence in the sample. Second, the optimal number of screening assays depends on probabilities of GMO presence in the sample; having more screening assays increases the screening cost but decreases the total expected cost if the screening assays are chosen smartly. Third, testing costs of *GMOtrack* strategies increase with the probability of GMO presence in the sample. In summary, both the current testing strategy costs and the *GMOtrack* testing strategy costs increase with the probability of GMO presence in the sample, but the *GMOtrack* costs are smaller and increase less than the costs of the current testing strategy.

Discussion

The idea we pursued through this study is that sample-centered GMO testing strategies that minimize the total GMO analysis cost can be automatically generated and bring major cost benefits. To prove this, we formalized the laboratory-level GMO traceability optimization problem and developed an algorithm, called *GMOtrack*, which systematically searches through possible strategies and generates cost-effective GMO traceability testing strategies. *GMOtrack* advises on the best set of screening elements to be applied in the first testing phase for a given sample without compromising the detection accuracy. Our experiments

Table 2. *GMOtrack* testing strategies through time[a]

	1996	1997	1999	2002	2005	2008	Beyond 2008
Number of GMOs	1	7	15	17	20	22	37
Probability of non-GMO sample	98	96	94	89	71	41	36
Number of screening assays	0	3	3	3	4	6	7+1 (8)
Optimal screening assays	RRS[b]	P-TA29	P-35S	P-35S	P-35S	P-FMV	CryIAb
		CP4 epsps	CTP2	CTP2	CTP2	CryIAb	nptII
		T-35S	bar	bar	bar	T-nos	pat
					pat	nptII	CP4 epsps
						P-35S::pat	manA (pmi)
						bar	P-35S::bar
							P-35S::pat
							LY038
Screening cost	0	155	155	155	177	219	261
Expected event-specific cost	113	6	15	30	58	82	108
GMOtrack cost	113	160	170	185	234	301	442
Current testing strategy cost		197	178	202	290	509[c]	741[c]

[a] Comparison of optimal testing strategies proposed by *GMOtrack* and the current testing strategy from year 1996 to 2008. The 1997 column represents the optimal testing strategy for data from Table 1. (*1*) Number of GMOs: Number of GMOs allowed in the EU. For "beyond 2008," the number of GMOs corresponds to the GMOs allowed in the EU plus the GMOs awaiting decision in April 2008. (*2*) Probability of non-GMO sample (%): the probability of a sample being non-GMO.

GMOs that were never detected were assigned a low frequency of one per thousand. For "beyond 2008," same frequencies of presence of each GMO were used as for the year 2008. GMOs awaiting decision were assigned a low frequency of one per thousand. (*3*) Number of screening assays: number of screening assays in the optimal testing strategy determined by the *GMOtrack*-algorithm. (*4*) Optimal screening assays: the optimal combination of screening assay determined by the *GMOtrack* algorithm. (*5*) Screening cost (EUR): Depends only on the number of screening assays according to the equation $g = 21.18 \cdot numAssays + 91.82$. (*6*) Expected event-specific cost (EUR): computed according to probabilities of screening outcomes and the number of event-specific assays needed in the second phase of the testing strategy. (*7*) *GMOtrack* cost (EUR): the sum of the screening cost and the expected event-specific cost as calculated by the *GMOtrack* algorithm. (*8*) Current testing strategy cost (EUR): The current testing strategy used by the case-study laboratory consists of a screening phase using the combination of two screening assays (P-35S, T-nos) and an event-specific assay (GT73), followed by a second phase (event-specific assays) depending on the outcome of the screening phase.

[b] For 1996, the optimal screening strategy would be event-specific RRS soybean assay only.

[c] From 2008 on, the current testing strategy does not allow detection of the presence of all authorized GMOs in the European Union.

Table 3. Optimal testing strategies depending on prior probabilities of GMO presence[a]

April 2008	Screening assays	Screening cost	Expected identification cost	Expected total cost	Optimal combination of screening assays
Food	3	155	78	233	P-35S, T-nos, P-FMV::epsps
	4	177	50	227	P-FMV, CrylAb, T-nos, T-35S
	5	198	32	230	P-FMV, CrylAb, T-nos, T-35S, nptII
	6	219	26	245	P-FMV, CrylAb, T-nos, nptII, pat, bar
	7	240	22	263	P-FMV, CrylAb, T-nos, nptII, pat, bar, CTP2
Feed	3	155	296	451	P-FMV, T-nos, P-35S
	4	177	193	370	P-FMV, CrylAb, T-nos, T-35S
	5	198	120	317	P-FMV, CrylAb, T-nos, nptII, P-35S::pat
	6	219	95	314	P-FMV, CrylAb, T-nos, nptII, P-35S::pat, bar
	7	240	83	323	P-FMV, CrylAb, T-nos, nptII, P-35S::pat, bar, CTP2

[a] Comparison of total expected costs of optimal combinations of different numbers of screening assays in two situations: given the food or feed prior probabilities of GMO presence in the sample. The total expected cost (in EUR) is the sum of screening and expected identification costs (in EUR); the numbers are rounded to increase readability.

showed that, by using these testing strategies, major savings can be achieved.

Usually, the laboratories involved in routine GMO diagnostics use a screening phase to preselect the samples that possibly contain GMOs before a final identification process using event-specific assays (4, 11). Screening relies on the use of the most common genetic elements found in GM crops, such as the cauliflower mosaic virus 35S promoter (P-35S) and the nopalin synthase terminator (T-nos; 12, 13). The choice of screening elements is intuitive and only provides information on the presence or absence of GMOs (11) without helping the final GMO identification. The results presented in this study show that, for a routine laboratory, the optimal screening combination is not the most intuitive P-35S and T-nos duplex assay (often completed with the nptII assay), but that it depends on the number and variety of GM crops needed to be detected (Figure 2) and with the probabilities of occurrence of these GM crops (Table 3). It also appears that, with the exception of 1996, when only the GM crop RRS was targeted, the P-35S and T-nos duplex assay is not able to cover all the targeted GM crops alone (Table 2).

Our TestLab currently uses a detection strategy with a screening phase combining the P-35S and T-nos screening assays and the RT73 event-specific assay. The results of the present study demonstrate the increasing cost per sample of this detection strategy, and that this combination is unable to ensure the detection of all authorized GM crops from 2008 on (Table 2). Moreover, it turns out that this combination was never the optimal one in terms of analysis costs (Figure 2). The cost ratio between the current detection strategy and the strategies proposed by *GMOtrack* constantly increases with years (Figure 2).

We also investigated the influence of prior probabilities of GMO presence on the optimal strategy. For this, we used the estimates of the frequency of appearance of GM crops in food and feed samples as gathered by our TestLab. Both types of samples present different possible occurrences of GMOs. From the results of this investigation (Table 3), it appears that the optimal testing strategy and the associated cost are affected by the prior probability of GMO presence in the sample, and vary between food and feed samples. This shows that, the more background knowledge the expert has on a sample to be tested, the more the combination of screening assays can be tuned to the sample. In our experiments, we set the probabilities based on past experience in routine analysis, but if additional information about GMO presence is known for a sample at hand, this should be encoded as probabilities, and *GMOtrack* would compute the optimal testing strategy for that sample. For these reasons, a unique optimal combination of screening assays does not exist for all types of samples.

Given the results of our investigation, using this new strategy would have an important impact in the daily life of an official GMO's control laboratory. By using the *GMOtrack* approach, routine laboratories will make significant savings in terms of experimental costs. Take as example a feed sample containing maize grains. In the case of our TestLab, it is known that such a maize sample is often contaminated with traces of soybean. The analyst would then follow the *GMOtrack* approach using a data table limited to relevant GMO events (maize and soybean). The computation in our TestLab shows that the best cost-efficient detection strategy uses a screening phase with the combination of three methods. This strategy is expected to save more than 45 euros/sample in comparison to the basic P-35S/T-nos screening, which, however, appears as the best two-component screening combination (Table 4). Similarly, in front of a packed food sample for which listed ingredients include maize, oilseed rape, and soybean, the analyst would find that the best screening strategy uses a combination of four methods

Table 4. Real case application on feed and food samples[a]

Use case	Sample composition	Screening assays	Screening cost	Expected identification cost	Expected total cost[b]	Optimal combination of screening assays
Feed	Maize grains and soybean dust	2	134	174	308	P-35S, T-nos
		3	155	106	262	CrylAb, T-nos, P-35S::pat[c]
		4	177	93	270	CrylAb, nptII, T-nos, P-35S::pat
		5	198	82	280	CTP2, CrylAb, nptII, T-nos, P-35S::pat
		6	219	81	299	CTP2, CrylAb, nptII, T-nos, Pat, T-35S
Food	Packed food: maize, oilseed rape, and soybean	3	155	65	220	P-35S, CTP2, T-nos
		4	177	40	217	CTP2, CrylAb, T-nos, T-35S[c]
		5	198	25	223	CTP2, CrylAb, Pat, Bar, T-nos
		6	219	22	241	CTP2, CrylAb, nptII, T-nos, P-35S::bar, P-35S::pat

[a] Total expected costs using combinations of different numbers of screening assays suggested by *GMOtrack* for two samples. Feed sample is composed of maize grains in which the analyst expects the presence of soybean dust traces. Food sample is packed food for which the list of ingredients includes maize, oilseed rape, and soybean. Computation was made using the food or feed prior probabilities of GMO presence in the sample.

[b] The total expected cost (in EUR) is the sum of screening and expected identification costs (in EUR); the numbers are rounded to increase readability.

[c] Most cost-efficient combination.

(including P-35S and T-nos). The basic P-35S/T-nos screening does not cover all authorized GMOs (Table 4).

It is still possible to intuitively choose the screening methods that are the most appropriate for screening all authorized GMOs in one's jurisdiction. However, with the increasing number of GMOs being commercialized, this intuitive setup will soon become impossible. Using the *GMOtrack* approach would therefore help routine detection in enforcement and companies' laboratories to ensure that all possible GM events authorized in the relevant jurisdiction are covered in an automated manner, thereby ensuring cost-efficient detection.

GMOtrack can be seen as the core algorithm of a future decision support system for routine use in laboratories. An engineering effort is necessary to provide a proper user interface with connections to external GMO data sources, support for GMO presence probability tuning, which might depend on sample type (e.g., food or feed sample), sample composition (e.g., prepared food with many ingredients or seed sample), sample origin, or even an automated sample history tracking. The system should also allow the user to provide other background information that can influence the optimal testing strategy, for example, the information about the reference gene systems for the sample. This system would then prepare the data table and feed it to the *GMOtrack* algorithm for generation of the optimal testing strategy, given the data and the parameters. The user interface should then also be able to interpret the wet-lab results and guide the user to the final identification of GMOs. Because the algorithm performs exhaustive search through an exponential search space, it will no longer be appropriate if data grow beyond some limit. In such a case, an algorithm using a heuristic approach will need to be developed. These further developments are planned for our future work.

Conclusions

The increasing presence of GMOs on the market and their growing taxonomic and biotechnological diversity are the main causes of increases in laboratory-level GMO testing costs. Our analysis of GMO traceability costs in time shows the adequacy of testing strategies currently used in routine laboratories just for the early years of GMO traceability. Their increased costs in the situation where dozens of GMOs need to be detected are alarming. Therefore, optimization approaches to analytic GMO traceability need to be developed.

Our research has the following implications. First, it proves that the currently used GMO testing strategy is unnecessarily expensive and may stimulate individual testing laboratories to change the approach, resulting in a shift from "the same strategy for all samples" to more cost-effective "sample-centered GMO testing strategies." Second, the *GMOtrack* algorithm can be used to theoretically evaluate new PCR assays; a theoretically founded evaluation of a PCR screening method with *GMOtrack* may prove or disprove the advantages in terms of cost of using such a method in routine practice. Third, the presented algorithm can be seen as the core element of a future decision support system to be used in routine laboratory practice. The implementation is flexible, allowing for changing of the data, probabilities, and parameters, so that what-if scenarios can be tested and

simulations of future testing costs can be computed depending on different parameters.

Testing strategies tailored to each sample could be generated if routine laboratories had decision support systems available for real-time strategy generation. This would substantially decrease the GMO testing costs, which would imply increased food safety and more control over GMO-related environmental issues.

GMOtrack was developed for GMO traceability, but its application can be extended to other domains where complex testing is involved, such as for safety and quality assurance in the food supply chain in general.

Acknowledgments

Authors Kralj Novak and Lavrač formalized the optimization problem, conceived the algorithm, and managed the information technology aspect. Authors Gruden, Morisset, Štebih, Rotter, and Žel conceived the study; acquired, preprocessed, and organized the data; and evaluated the results. We all contributed to, read, and approved the final manuscript.

We acknowledge Lotte Hougs (National Institute of Biology, Ljubljana, Slovenia), who provided part of the data present in the GMO database used in this work, and Jana Vojvoda (Danish Plant Directorate, Lyngvy, Denmark) who contributed with the data preparation.

Funding was provided by Slovenian Research Agency (P4-0165, P2-0103); Ministry of Agriculture, Forestry and Food; and Ministry of the Environment and Spatial Planning (V4-0314).

References

(1) James, C. (2009) *Global Status of Commercialized Biotech/GM Crops: 2008*, International Service for the Acquisition of Agri-Biotech Applications, Ithaca, NY

(2) Gruere, G.P., & Rao, S.R. (2007) *AgBioForum* **10**, 51–64

(3) Miraglia, M., Berdal, K.G., Brera, C., Corbisier, P., Holst-Jensen, A., Kok, E.J., Marvin, H.J., Schimmel, H., Rentsch, J., van Rie, J.P., & Zagon, J. (2004) *Food Chem. Toxicol.* **42**, 1157–1180

(4) Holst-Jensen, A. (2007) in *Food Toxicants Analysis: Techniques, Strategies and Developments*, Y. Pico (Ed.), Elsevier, Amsterdam, The Netherlands, pp 231–268

(5) Agbios, GM database, http://www.agbios.com/dbase.php

(6) European Commission, EU community register of authorized GM food and feed, http://ec.europa.eu/food/dyna/gm_register/index_en.cfm

(7) GMO compass, GMO database, http://www.gmo-compass.org/eng/gmo/db/

(8) Organisation for Economic Co-operation and Development, Biotrack Product Database, http://www2.oecd.org/biotech/

(9) Biosafety Clearing House, Living Modified Organism (LMO) Registry, http://bch.cbd.int/database/organisms/uniqueidentifiers/

(10) Bruderer, S., & Leitner, K.E. (2003) *Genetically Modified (GM) Crops: Molecular and Regulatory Details*, Centre for Biosafety and Sustainability, Zurich, Switzerland

(11) Aarts, H.J.M., Rie, J.P.P.F., & Kok, E.J. (2002) *Expert Rev. Mol. Diagn.* **2**, 69–77

(12) Cardarelli, P., Branquinho, M.R., Ferreira, R.T.B., da Cruz, F.P., & Gemal, A.L. (2005) *Food Control* **16**, 859–866

(13) Vollenhofer, S., Burg, K., Schmidt, J., & Kroath, H. (1999) *J. Agric. Food Chem.* **47**, 5038–5043

Supplemental Information

Supplemental information is available online at http://www.atypon-link.com/AOAC/loi/jaoi including sections on *Formal Problem Definition, GMO Traceability as an Optimization Problem*, and *The GMOtrack Algorithm*.

The *GMOtrack* Web page http://kt.ijs.si/software/GMOtrack/ provides (*1*) the data that was used in the experiments in this paper, (*2*) the implementation of the *GMOtrack* algorithm (executable and source code), (*3*) the real-time PCR analysis costs schema, and (*4*) instructions and other related information.