

SCIENTIFIC REPORTS



OPEN

Metabolomic approach for a rapid identification of natural products with cytotoxic activity against human colorectal cancer cells

Vittoria Graziani¹, Monica Scognamiglio^{1,2}, Valentina Belli³, Assunta Esposito¹, Brigida D'Abrosca¹, Angela Chambery¹, Rosita Russo¹, Marta Panella¹, Aniello Russo ¹, Fortunato Ciardiello³, Teresa Troiani³, Nicoletta Potenza¹ & Antonio Fiorentino¹

The discovery of bioactive compounds from natural sources entails an extremely lengthy process due to the timescale and complexity of traditional methodologies. In our study, we used a rapid NMR based metabolomic approach as tool to identify secondary metabolites with anti-proliferative activity against a panel of human colorectal cancer cell lines with different mutation profiles. For this purpose, fourteen Fabaceae species of Mediterranean vegetation were investigated using a double screening method: ¹H NMR profiling enabled the identification of the main compounds present in the mixtures, whilst parallel biological assays allowed the selection of two plant extracts based on their strong anti-proliferative properties. Using high-resolution 2D NMR spectroscopy, putative active constituents were identified in the mixture and isolated by performing a bio-guided fractionation of the selected plant extracts. As a result, we found two active principles: a cycloartane glycoside and protodioscin derivative. Interestingly, these metabolites displayed a preferential anti-proliferative effect on colon cancer cell lines with an intrinsic resistance to anti-EGFR therapies. Our work provides an NMR-based metabolomic approach as a powerful and efficient tool to discover natural products with anticancer activities circumventing time-consuming procedures.

The development of plant-derived substances has established the basis of sophisticated traditional medicine and, nowadays, plants still continue to be an essential source of pharmaceutical agents^{1,2}. In particular, natural products have been the pillar of cancer chemotherapy for several years. Even though combinatorial chemistry provides a wide range of new and synthetic drugs, natural products are likely to provide lead compounds for the development of novel agents with enhanced biological properties^{3,4}.

This success in drug discovery is linked to the high chemical diversity of natural sources; nevertheless, the chemical variability and the vast number of metabolites, which occur in the plant kingdom, make the detection and characterisation processes challenging to perform⁵. Traditional methodologies to discover new active molecules have different drawbacks, such as their complexity and inherent lengthy timescales, given that this research requires different separation and isolation steps before an active compound will emerge from the crude extract. Recently, the screening of plant-derived mixtures has become an effective way for rapid selection of metabolites with biological properties^{6,7}.

In this scenario, metabolomics can provide a pertinent and convenient strategy, by allowing the direct study of the crude extract, simultaneously observing a wide range of metabolites belonging to different chemical classes, without the need of time-consuming isolation procedures⁸. The rapid sample preparation, shorter NMR measurement times, advanced data analysis methods and the use of advanced two-dimensional (2D)

¹Dipartimento di Scienze e Tecnologie Ambientali Biologiche e Farmaceutiche (DiSTABiF), Università degli Studi della Campania "Luigi Vanvitelli", via Vivaldi 43 I-, 81100, Caserta, Italy. ²Max Planck Institute for Chemical Ecology - Beutenberg Campus, Hans-Knöll-Straße 8 D-, 07745, Jena, Germany. ³Dipartimento di Internistica Clinica e Sperimentale "Flaviano Magrassi", Università degli Studi della Campania "Luigi Vanvitelli" Via Pansini, 5 -, I-80131, Napoli, Italy. Vittoria Graziani and Monica Scognamiglio contributed equally to this work. Correspondence and requests for materials should be addressed to T.T. (email: teresa.troiani@unicampania.it) or N.P. (email: nicoletta.potenza@unicampania.it) or A.F. (email: antonio.fiorentino@unicampania.it)

techniques to determine the structures of known or unknown compounds, make the NMR-based metabolomic approach extremely powerful for the profiling of crude extracts and the rapid identification of natural products⁹. Furthermore, relating putative biological activities with metabolomics is feasible by integrating chemometric methods with bioactivity results^{10,11}. In this work, we combined an NMR-based metabolomic approach with MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) tetrazolium salt colorimetric assays to select plant species belonging to Fabaceae family with anti-proliferative properties against a panel of genetically different colon cancer cell lines.

Colorectal cancer is one of the most frequently diagnosed malignant diseases in Europe and one of the leading causes of cancer-related death worldwide¹². Although the outcome of patients with metastatic colorectal cancer (mCRC) has improved during the last years, current therapies available are not sufficiently effective¹³ and their efficacy is limited by the side effects of the drugs and/or the development of resistance^{14,15}. For these reasons, there is a resurgence of interest for natural products, such as medicinal plants and dietary means as an alternative solution to cure mCRC patients. In the plant kingdom, the Fabaceae family represents a heritage of high biodiversity for their richness in secondary metabolites¹⁶, particularly significant for human health, either as dietary supplements or as pharmacological agents¹⁷. Furthermore, many studies reveal an interesting correlation between Fabaceae activities and colon cancer prevention and therapy, identifying this family as a promising source of new effective molecules against colon cancer. Soy (*Glycine max*) contains many bioactive compounds, including isoflavones. Epidemiological studies have shown that a high-level intake of soy-derived products contributes to a lower incidence of colorectal cancer in Asian countries. Among various soy isoflavones, genistein (4,5,7-trihydroxyisoflavone) (GEN) has been evaluated as a good candidate in colon cancer prevention¹⁸. The formononetin, constitutive isoflavones from barrel medic (*Medicago truncatula*), exerts strong antiangiogenic effects reducing colon cancer proliferation both *in vitro* and *in vivo* experiments¹⁹. A naturally occurring rotenoid deguelin, isolated from *Mondulea sericea*, an African legume species, significantly inhibited IL-8 gene expression, induces apoptosis in colon cancer cells by down regulating the NF- κ B signalling²⁰. Isoliquiritigenin, a chalcone from roots of liquorice roots (*Glycyrrhiza glabra*) could be a promising adjuvant in colon cancer treatment, reducing tumour growth and protecting kidney and liver against chemotherapy-induced toxicity in a mouse xenograft model of colon carcinoma²¹. The triterpenoid B-group soyasaponins have been found to induce macroautophagy in human colon cancer cells²².

Thus, in the present work we applied a new screening method to seek secondary metabolites from fourteen Mediterranean Fabaceae species, which could be promising lead-structures to develop novel agents in CRC chemotherapy.

Results

Metabolic profiling of the selected Fabaceae species. The selected plants (*Astragalus boeticus* L., *Lathyrus cicera* L., *Lathyrus clymenum* L., *Medicago minima* (L.) L., *Melilotus neapolitanus* Ten., *Ononis variegata* L., *Pisum sativum* L. subsp. *biflorum* (Raf.) Soldano, *Trifolium campestre* Schreb., *Trifolium cherleri* L., *Trifolium scabrum* L. subsp. *scabrum*, *Trigonella esculenta* Willd., *Vicia bithynica* (L.) L., *Vicia pseudocracca* Bertol., *Vicia angustifolia* L.) were analysed for their metabolic phenotype by multinuclear NMR spectroscopy. ¹H NMR profiling of these plant extracts allowed the identification of the main compounds present in the mixture. The metabolite assignment was done by comparing peak chemical shifts to those found in literature²³ and in Human Metabolome Database (HMDB), and furthermore, all the metabolite structures were confirmed by 2D NMR experiments. The peculiar NMR values of 31 identified metabolites were reported in Table S1. This analysis revealed that all the studied species had a similar composition in primary metabolites (amino acids, organic acids and sugars) but a high variability was seen in terms of secondary metabolites, the majority of which were phenols. A group of compounds well represented in all the analysed species are flavonoids, characterised by distinctive doublets in the aromatic region of the proton spectra. Among these, quercetin, kaempferol and their derivatives were the most common compounds. Other flavonoids, on the contrary, were only present in selected species: a C6-glycosylated apigenin derivative and catechins were found in *T. campestre* and *L. clymenum*, respectively; meanwhile, in the *M. minima* ¹H-spectrum different doublets in the region between 6.30 and 7.50 ppm supported the presence of isoflavone compounds such as daidzein, daidzin and genistein.

Cinnamic acid derivatives and caffeic acid were also widespread in all the studied plants. Chlorogenic acids were found in *V. bithynica*, *V. angustifolia*, while coumarin and their glucoside precursors such as *cis/trans*-melilotoside and dihydromelilotoside were detected in *M. neapolitanus*²⁴. At lower fields, the singlet at δ_{H} 9.15 and the triplet at δ_{H} 8.86 indicated the presence of trigonelline, a methylbetaine derivative of nicotinic acid. In the opposite region of spectra, signals of terpene metabolites were detected in *A. boeticus*. In detail, two doublets at δ_{H} 0.37 (J = 4.6 Hz) and δ_{H} 0.57 (J = 4.6 Hz) revealed the presence of cycloartane compounds in this plant extract.

Cytotoxic activity of plant extracts. The potential anti-proliferative activity of the plant extracts studied here was evaluated against three human colorectal cancer cell lines (Caco-2, HT-29 and HCT-116), which were selected for their diverse mutation profiles and their resulting different medical aspects. In particular, Caco-2 cell line has no genetic alterations in KRAS, NRAS, BRAF and PIK3CA genes, which were known to be associated with intrinsic resistance to anti-EGFR therapies, such as cetuximab and panitumumab. On the contrary, HT-29 and HCT-116 cells displayed a different genetic profile, harbouring BRAF V600E and KRAS/PIK3CA, respectively.

Cancer cells were treated with increasing doses of plant extracts ranging between 10 to 250 $\mu\text{g/ml}$ for 48 hours and their potential cytotoxicity was assessed through MTT assays. The resulting data expressed as percentage values of the cell growth respect to the control were used to develop multivariate statistical analyses, in particular the agglomerative hierarchical clustering (HCA) and the principal component analysis (PCA), which can provide

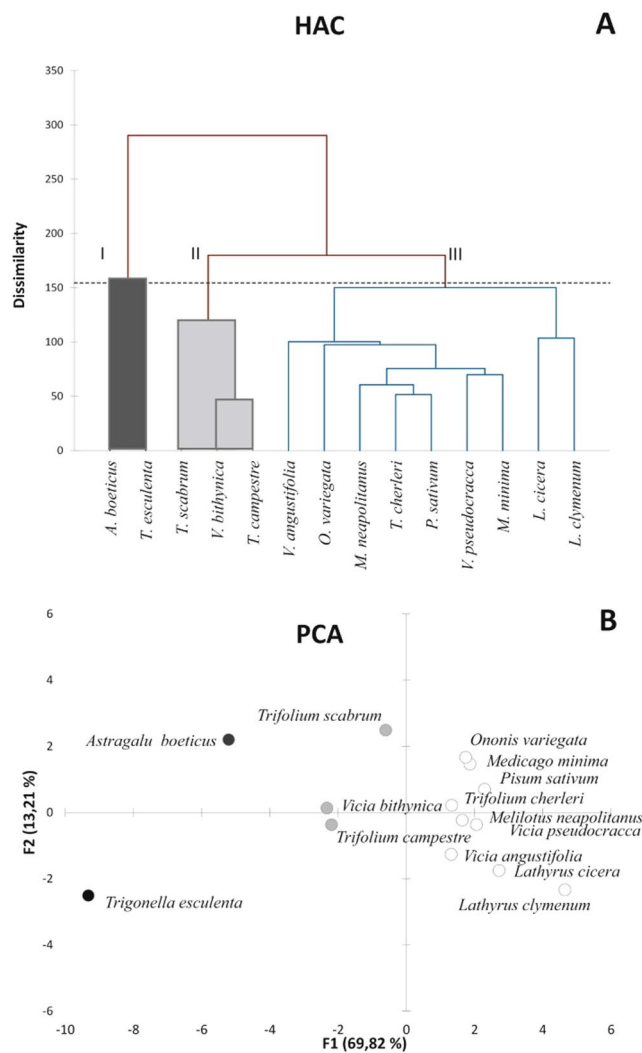


Figure 1. HCA dendrogram (A) and PCA (B) of cell growth percentage from control of colon cancer cell lines treated with the selected plant extracts over 48 hours. Based on these analyses the species are classified in three subsets. Groups I included the active species (black), II those active only at the highest tested doses (grey) and III those that have no-significant effect (white).

an overview of the results as well as a categorisation of the experimental dataset. Indeed, the HCA dendrogram clearly pointed out three different clusters with a high dissimilarity degree (150%). The first group included all treatments with *A. boeticus* and *T. esculenta* plant extracts, the second was formed by *T. scabrum*, *V. bithynica* and *T. campestre*, while the third cluster encompassed all of the remaining plants (Fig. 1A). This trend was further confirmed by PCA (Fig. 1B), where in the score scatter plot the three plant groups were mainly arranged along a gradient generated by the first component, which accounted for 72.84% of the variance. This, therefore, indicated an extreme diverse anti-proliferative activity amongst the investigated plant extracts pointing out that the maximum effect on the cell growth was caused by the above-mentioned first group. In particular, *T. esculenta* showed a similar strong anti-proliferative effect on the Caco-2, HT-29 and HCT-116 cell lines, whilst *A. boeticus* displayed a mild activity on Caco-2 and HCT-116 cell lines but a selective and notable anti-proliferative effect on HT-29 cell line. These findings suggested that *A. boeticus* extract could operate with a selective mode of action on the HT-29 colon cancer cells.

The second group identified in the PCA included plants that showed a severe anti-proliferative effect only at the highest tested concentrations (200 and 250 µg/ml). The rest of the screened plants (the third group) did not exhibit cytotoxic activity under the experimental conditions used and in some cases (*Lathyrus* spp vs. HT-29 cell lines) they exerted a stimulating effect on the cell growth. Based on these results, an extensive 2D NMR analysis was performed on the active plant extracts (*A. boeticus* and *T. esculenta*), with the aim of identifying the compounds in the mixtures that were responsible for the observed cytotoxicity (Fig. 2).

2D NMR analysis of the active plant extracts. *Astragalus boeticus.* The NMR analysis of *A. boeticus* extract showed the presence of three main secondary metabolites. The first two molecules were identified as

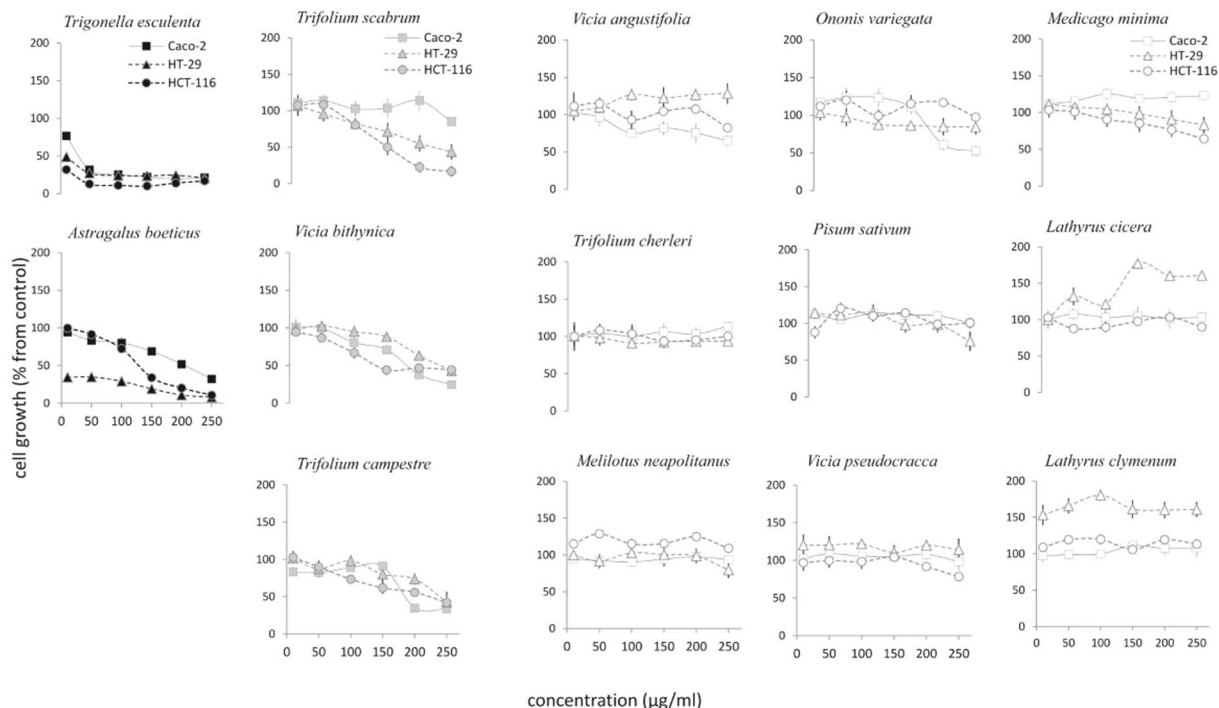


Figure 2. Cytotoxicity of the studied Fabaceae species on Caco-2, HT-29 and HCT-116 colorectal cancer cell lines evaluated by MTT assays over 48 hours. Cell growth is expressed as percentage from control and it is plotted in the vertical scale, while doses of plant extracts are depicted in the horizontal axis. Groups I, II and III are depicted in black; grey and white, respectively.

quercetin and kaempferol (Table S1), two flavonols widely distributed in the plant kingdom, while the third metabolite was a cycloartane triterpene. Beyond the methylene H-19 protons at δ_{H} 0.37 and 0.57, seven methyl singlets were clearly detectable in the ^1H NMR spectrum (Fig. 3A). In detail, signals at δ_{H} 1.01 (δ_{C} 15.5) and 1.06 (δ_{C} 25.7) were assigned to the H-28 and H-29 methyls, based on HMBC heterocorrelations (Fig. 4A) with C-5 methine (δ_{C} 49.5), carbon that was found to bind the proton at δ_{H} 1.66. This proton (H-5) correlated with the H-19 methylene and H-6 proton in the COSY experiment; this resonated at δ_{H} 4.75 suggesting the presence of an oxygen atom binding the C-6 carbon. Furthermore, both C-28 and C-29 carbons correlated with the H-5 methine and H-3 proton at δ_{H} 4.44 (δ_{C} 88.4). Besides, the H-19 protons displayed heterocorrelations with the C-8 methine carbon (δ_{C} 45.0), which in turn correlates with the H-30 methyl that resonated at δ_{H} 0.99 (δ_{C} 19.2). Furthermore, the signal at δ_{H} 1.24 (δ_{C} 20.3), attributed to the 18 methyl, and the peak at δ_{H} 1.26 (δ_{C} 27.2), related to the H-21 methyl, correlated with the C-17 carbon at δ_{C} 57.7. In the COSY spectrum, the proton bound to this carbon revealed homocorrelations with the H-16 proton at δ_{H} 4.64; this chemical shift value was well in agreement with the presence of a hydroxyl group in this position. A series of HMBC correlations, which were found between the H-21 proton and the C-20 carbon (δ_{C} 87.6); the H-26 methyl (δ_{H} 1.24) and the C-25 carbinol (δ_{C} 72.0); the H-27 methyl (δ_{H} 1.34) and the C-24 carbinol (δ_{C} 81.2) allowed us to hypothesise the presence of a substituted tetrahydrofuran moiety in the side chain of the triterpene. All these data were in agreement with the presence of a cycloastragenol^{25,26}. The downfield shift of the H-3 and C-3 values justified the presence of a sugar moiety, indicating a presumable site of glycosylation at position 3. This hypothesis was validated with an HMBC experiment, in which the C-3 carbon clearly heterocorrelated with the anomeric proton at δ_{H} 4.44 (δ_{C} 107.0). Finally, the H-6 proton showed cross peak with a carbonyl at δ_{C} 172.5, which in turn correlated with the methyl at δ_{H} 2.04 (δ_{C} 21.2). These data demonstrated the presence of an acetyl group, which formed an ester bond with the hydroxyl group located at C-6 carbon.

Trigonella esculenta. *T. esculenta* metabolome presented little variability in terms of secondary metabolites; indeed, the ^1H spectrum was dominated by 4-hydroxyisoleucine signals (Fig. 3B). In an attempt to see beyond these primary metabolites, the crude extract was purified on a Sep-pak C-18 cartridge and the resulting methanol eluate was analysed by NMR. In the up-field region of the spectrum, peculiar doublets and singlets between 4.2 and 5.2 ppm suggested the presence of triterpenoid saponins. In particular, two singlets at δ_{H} 0.83 and 1.04, alongside two doublets at δ_{H} 0.94 and 1.01 suggested the presence of a steroid moiety²⁷. In the HMBC experiment (Fig. 4B), the doublet at δ_{H} 0.94 (δ_{C} 17.6) correlated either with the methine C-24 at δ_{C} 35.2 or a methylene carbinol C-28 at δ_{C} 75.9 (δ_{H} 3.70 and 3.40). This signal showed cross peaks with the doublet at δ_{H} 4.73 (δ_{C} 104.8), indicating a glycosylation site at position 28. Moreover, the doublet at δ_{H} 1.01 (δ_{C} 15.7) heterocorrelated with two methines at δ_{C} 40.8 and 64.6, and with the acetal carbon at δ_{C} 113.9. These data are consistent with the presence of a furostanol moiety in the molecule. A second site of glycosylation was identified at position 3; indeed, the

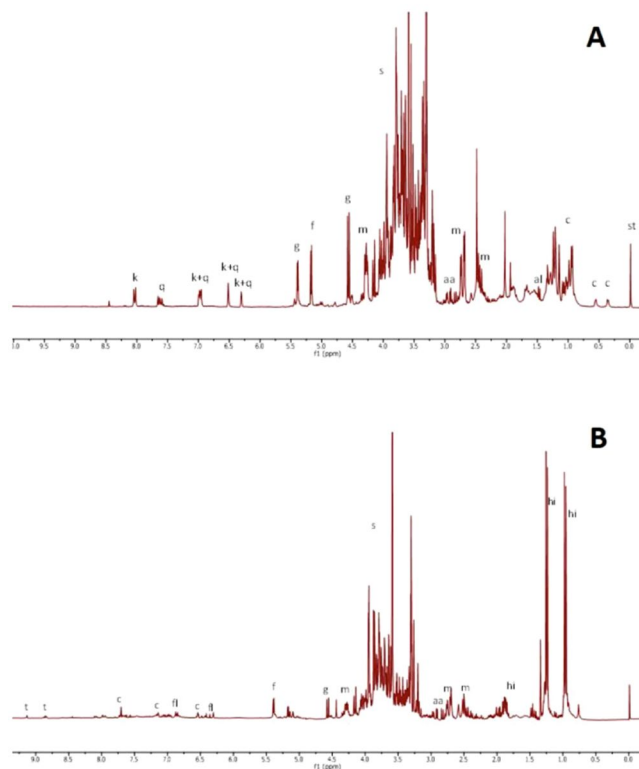


Figure 3. (A) ^1H NMR of *Astragalus boeticus* extract registered in phosphate buffer in D_2O (pH 6.0) and methanol d_4 (1:1). (B). ^1H NMR of *Trigonella esculenta* registered in phosphate buffer in D_2O (pH 6.0) and methanol d_4 (1:1). ABBREVIATIONS: aa, aspartic acid; al, alanine; c, cycloartane; fl, flavonoids; f, fructose; g, glucose; h, 4-hydroxyisoleucine; k, kaempferol; malic acid; q, quercetin; st, standard; s, sugar; t, trigonelline.

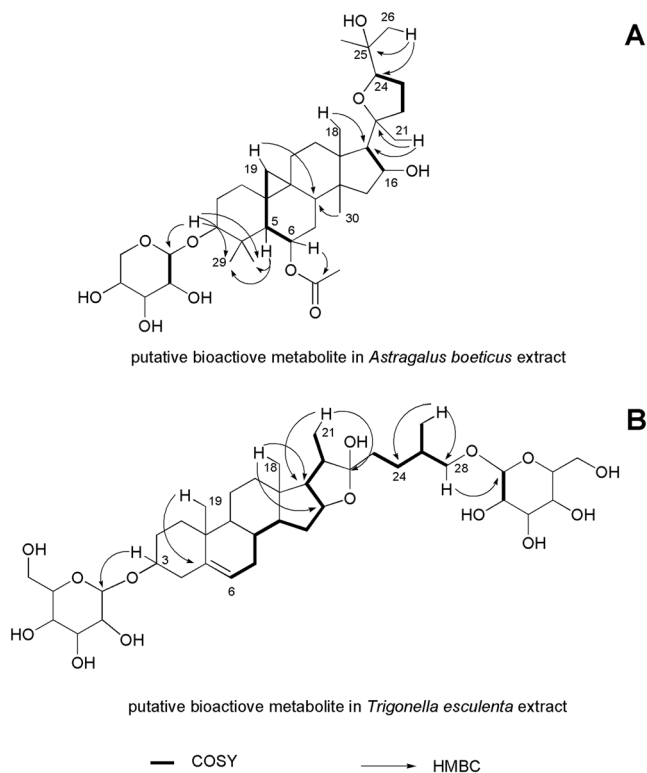


Figure 4. Main COSY and HMBC correlations of the putative bioactive metabolites in *Astragalus boeticus* and *Trigonella esculenta* extracts.

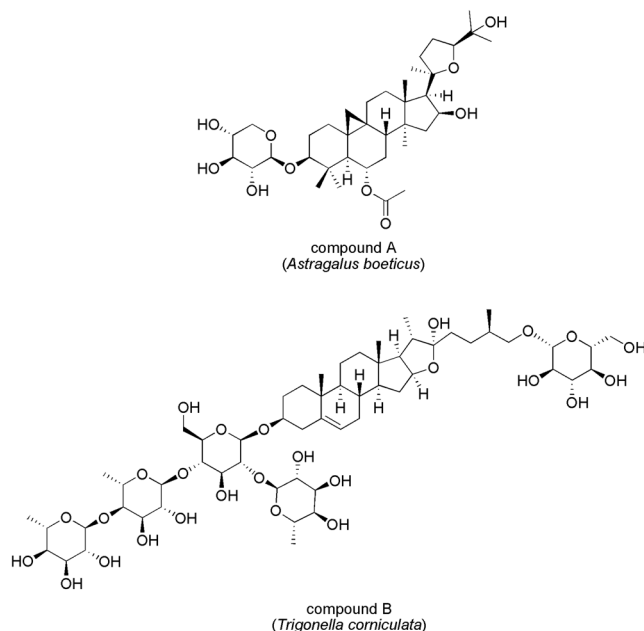


Figure 5. Bioactive constituents of *Astragalus boeticus* (compound A) and *Trigonella esculenta* (compound B) extracts.

carbinol carbon at δ_C 79.2 correlated with the anomeric proton at δ_H 4.40. Moreover, overlapped doublets between 1.20 and 1.30 ppm correlated in the HSQC experiment with carbons at δ_C 18.4, 18.1 and 17.9. Finally, the correlations of these protons with carbinol carbons supported the presence of three deoxy sugars.

Isolation and identification of the active metabolites from the cytotoxic plant extracts.

Compound A from *Astragalus boeticus*. *A. boeticus* crude extract was partitioned between ethyl acetate and water; after purification through different chromatographic processes, the organic phase led to the isolation of the pure compound A²⁸ (Fig. 5). It showed a molecular formula $C_{37}H_{60}O_{10}$, according to the ^{13}C NMR data and the positive ESI Q-TOF HRMS spectrum that displayed a sodiated adduct at m/z 687.4480. In the 1H NMR spectrum, characteristic signals of a cycloartenane triterpene were detected: two methines at δ_H 0.61 and 0.40 (H-19) and seven singlet methyls at δ_H 1.27, 1.26, 1.22, 1.13, 1.09, 1.01, 0.99. The presence of a doublet at δ_H 4.26, as well as other protons resonating between 3.18 and 3.46 ppm, suggested the presence of sugar in the molecule. The 2D NMR experiments (COSY, TOCSY, NOESY, HSQC, CIGAR-HMBC, H2BC and HSQCTOCSY) were in agreement with the presence of a cycloartenol as aglycone. In particular, in the CIGAR-HMBC experiment the H-6 proton at δ_H 4.75 showed a cross peak with the ester carbonyl at δ_C 171.7, which in turn correlated with the methyl at δ_C 1.99. These data suggested the presence of an acetate group at position 6. Moreover, a glycosylation site at position 3 was supported by the correlation between H-3 methine at δ_H 3.21 and the anomeric carbon at δ_C 107.4. The TOCSY, HSQCTOCSY and H2BC NMR experiments allowed the identification of the saccharide moiety as xylose, while the coupling constant value of the anomeric proton (6.0 Hz) agreed with a β configuration. Finally, the stereochemistry of compound A was defined using a NOESY experiment. Taken all together, these data enabled the identification of compound A as 6-O-acetyl-3-O- β -D-xylopiranosylcycloastragenol. This was confirmed by tandem MS analysis of the pseudomolecular ion $[M + Na]^+$ that showed the presence of the fragment ion at m/z 537.3399 and 627.3801, due to the loss of the xylose and the acetate, respectively.

Compound B from *Trigonella esculenta*. The crude extract of *T. esculenta* was chromatographed on Amberlite XAD-4 and eluted with water to eliminate polar metabolites. The saponin fraction, recovered by methanol, was fractionated through column chromatography and finally purified by RP-HPLC to obtain pure compound B^{29,30} (Fig. 5). It showed a molecular formula $C_{57}H_{94}O_{26}$, according to the positive ESI Q-TOF HRMS analysis and ^{13}C NMR data. The 1H NMR spectrum showed the signals of the steroidal aglycone as two singlet methyls at δ_H 0.84 and 1.05, and two doublets at δ_H 0.95 and 1.01.

In the 1H NMR spectrum, different anomeric protons suggested the presence of an oligosaccharide, which was located at C-3 carbon, thanks to the CIGAR-HMBC experiment. The HSQCTOCSY and H2BC experiment in combination with the Q-TOF HRMS analysis further confirmed this structure. Thus, compound B was unequivocally identified as [(25R)-furost-5-ene-3 β ,22 α ,26-triol 3-O- α -L-rhamnopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranosyl-(1 \rightarrow 4)-[α -L-rhamnopyranosyl-(1 \rightarrow 2)]- β -D-glucopyranosyl 26-O- β -D-glucopyranoside.

Anti-proliferative activity of pure compounds. In an attempt to understand whether the isolated compounds were actually responsible for the cytotoxicity of the crude extracts, their cytotoxicity was evaluated against a panel of colorectal cancer cell lines (Caco-2, HT-29, HCT-116). In these assays, a wide range of treatment concentrations was examined over 48 hours.

Active metabolites	Caco-2	HT-29	HCT-116
Compound A	50 μ M	3 μ M	40 μ M
Compound B	3 μ M	3 μ M	2 μ M

Table 1. IC₅₀ values of compound A and B evaluated by MTT assays over 48 hours on Caco-2, HT-29, HCT-116 colorectal cancer cell lines.

With regard to *A. boeticus*, after incubating the cells with compound A at concentrations below 50 μ M, a reduction of approximately 50% of cell growth was observed (Table 1). This strong antiproliferative effect resembled closely what had been observed with the crude plant extract. Interestingly, this metabolite did not exert the same anti-proliferative effect on the three types of colon cancer cells tested, as its IC₅₀ value was significantly lower for the HT-29 cells (IC₅₀ = 3 μ M). This may suggest a selective mode of action of compound A on this particular cell type.

With regard to *T. esculenta*, compound B was able to recapitulate the cytotoxic activity of the crude plant extract (Table 1). After the incubation of Caco-2, HT-29 and HCT-116 cell lines, a strong anti-proliferative effect was detected and very low doses (<10 μ M) reduced cell growth by approximately 50%, indicating that this metabolite was largely accountable for the anti-proliferative effect of the plant extract. Compound B displayed the same effect on Caco-2 and HT-29 (IC₅₀ = 3 μ M), while its activity was slightly higher in HCT-116 cell lines (IC₅₀ = 2 μ M).

To further validate that A and B were the active constituents of *A. boeticus* and *T. esculenta*, the other main secondary metabolites, which were previously identified in these plant derived-mixtures, were isolated and tested. In detail, we purified the flavonols quercetin and kaempferol from the first plant extract, and the amino acid 4-hydroxyisoleucine from the second. We subsequently evaluated their cytotoxicity on the selected human colon cancer cell lines, demonstrating that these have no significant effect under the examined conditions.

Discussion

Natural product research requires new strategies to renovate the traditional methodologies, which are too expensive and/or time consuming^{31,32}. Metabolomics ameliorate the effectiveness of previous approaches restoring the relevance of natural products as source of novel potential anticancer agents^{33,34}. Here, we integrate metabolomic strategies with biological assays to search secondary metabolites from Fabaceae species, which exert cytotoxicity against colon cancer cells.

We performed a detailed metabolomic analysis of fourteen Fabaceae species, unambiguously identifying 31 metabolites from these plants. With regard to the non-active species, there is a fragmented literature that described these from different points of view^{24,35–38}. Nonetheless, the current research unveiled a more comprehensive picture of the studied species, showing their metabolomic profiles and pointing out the main secondary metabolites present in the mixture. Because of their strong cytotoxic effect on Caco-2, HT-29 and HCT-116 human colorectal cancer cell lines, *A. boeticus* and *T. esculenta* extracts were selected for further investigations. To our knowledge, there is a little available data about these selected plant extracts; however, previous published work on other species belonging to the same genus assisted us in the search and identification of the putative active constituents of our selected plants.

In Traditional Chinese Medicine (TCM) *Astragalus* genus is particularly widespread for its antiperspirant, diuretic, and tonic properties. In particular *A. membranaceus*, is very well-known in folk medicine for its anti-inflammatory and immunomodulatory effects mainly exerted by regulating the (NF)- κ B signalling pathways^{39–41}. Furthermore, recent studies have proven that cycloartane and oleanan-type saponins modulate the cytokine release and might contribute either to the anti-inflammatory or anti-proliferative effect of *Astragalus* spp.⁴². It has been also demonstrated that *Astragalus* saponins from *A. membranaceus* regulate cell invasiveness and angiogenesis in human gastric adenocarcinoma cells⁴³.

Trigonella esculenta has been already investigated for its ethyl-galactoside⁴⁴ and diosgenin contents⁴⁵, yet no previous studies evaluated the anti-proliferative activity of this species. The majority of literature data refers to *Trigonella foenum-graecum* L. (Fenugreek), a medicinal and dietary plant to which were attributed a very wide range of biological properties⁴⁶, especially those regarding to the hypocholesterolaemic and hypoglycaemic effect^{47,48}. Notably, other previous works reported the potentiality of *T. foenum-graecum* as source of potential anti-cancer agents: in detail, anti-proliferative metabolites have been identified to be steroidal furostan-type saponins, such as diosgenin and protodioscin^{49,50}.

In this work, we sought to identify the compounds responsible for plant extract cytotoxicity against colorectal cancer cells, and to this aim an extensive 2D NMR analysis of the active crude extracts was carried out followed by a bio-guided fractionation to isolate these metabolites. Consistently with the previous investigations, we found a cycloartane glycoside (compound A) and a protodioscin derivative (compound B) as active principles from *A. boeticus* and *T. esculenta*, respectively (Fig. 5).

A previous work described compound A and its capacity to inhibit nitric oxide production, identifying this metabolite as a potential anti-inflammatory agent²⁷. Nevertheless, here the proliferation reducing activity of compound A was demonstrated for the first time. Supporting our findings, antineoplastic properties have already been associated with other cycloartane glycosides. Of these, astragaloside I – II – III isolated from *A. membranaceus* have been tested against a panel of different human cancer cells, which represent diverse common malignancies. The strongest anti-proliferative and apoptotic effect was observed in the leukemic (HL-60, HL-60/Dox and SKW-3) and in breast cancer (MDA-MB-231) cell lines, while the colorectal carcinoma HT-29 was found to be the least responsive to the saponin treatment⁵¹. This latter finding is not in full agreement with our results, where

compound A was found to be most active in HT-29 cells ($IC_{50} = 3 \mu M$) compared with the other studied colon cancer cell lines, Caco-2 ($IC_{50} = 50 \mu M$) and HCT-116 ($IC_{50} = 40 \mu M$). Another study described astragaloside II as a novel promising agent for reversal of multidrug resistance (MDR) mediated by P-glycoprotein in human hepatic cancer cells⁵².

Our work established for the first time the presence of compound B in *Trigonella spp.*, albeit this metabolite has already been isolated from other genus. Previous evidence proved that fractions deriving from *Livistona chinensis* var *subglobosa* seeds containing compound B promote apoptosis in different human cancer cells and, thus, it could be used to both prevent and treat cancer (US 2010/0179098 A1). Another invention patent included this molecule in a pharmacological composition together with a histone deacetylase inhibitor, the combination of these compounds causes an increase of the apoptotic rate (US 2012/0058961 A1). Here, we found that compound B has a potent anti-proliferative effect towards the investigated panel of colon cancer cell lines, with a slight preferentiality for HCT-116.

Interestingly, compound A and B displayed an enhanced effect on HT-29 and HCT-116 cells, respectively. These cell lines are characterised by genetic defects in BRAF and KRAS/PIK3CA, respectively. KRAS mutation, or genetic variation in other intracellular downstream effectors of EGFR activation, such as BRAF, NRAS, and PIK3CA, are responsible for the intrinsic resistance to anti-EGFR therapies^{53,54}. Moreover, the remarkable selectivity of compound A for HT-29 may have a potent clinical impact; since BRAF mutated mCRC patients have a very negative prognosis. In fact, for this pathology at present there is neither a standard chemotherapeutic treatment nor targeted therapies available. Although inhibition of BRAF oncoprotein by the small-molecule drug PLX4032 (vemurafenib) is highly effective in the treatment of melanoma, colon cancer patients associated with the same BRAF mutation have poor prognosis showing only a very limited response to this drug. As Vemurafenib treatment induces EGFR feedback activation, this may explain the refractoriness of BRAF (V600E) mutated colon cancers to this therapy⁵⁵.

In conclusion, using an innovative and highly efficient screening approach we quickly identified two natural products (A and B) as potent cytotoxic molecules against human colorectal cancer cells, which exert resistance against anti-EGFR therapies. These findings encouraged further investigations to understand the mode of action of compounds A and B and their potential application in drug resistant colon cancer therapy. Obviously, different limitations are intrinsically associated to our *in vitro* experimental system; and for this reason, future experiments will be also addressed to evaluate the pharmacokinetic properties and the bioavailability of the active metabolites in animal models.

Materials and Methods

Plant material. Plant species were harvested in April 2014 at the “Castel Volturno” Nature Reserve (40°57.587'N, 14°00.105'E; southern Italy), identified and the voucher specimens were deposited at the Herbarium of DiSTABiF of Università degli Studi della Campania “Luigi Vanvitelli, as reported below: *A. boeoticus* L., *Lathyrus cicera* L., *Lathyrus clymenum* L., *Medicago minima* (L.) L., *Melilotus neapolitanus* Ten., *Ononis variegata* L., *Pisum sativum* L. subsp. *biflorum* (Raf.) Soldano, *Trifolium campestre* Schreb., *Trifolium cherleri* L., *Trifolium scabrum* L. subsp. *scabrum*, *Trigonella esculenta* Willd., *Vicia bhytinica* (L.) L., *Vicia pseudocracca* Bertol., *Vicia angustifolia* L.. Leaf samples were collected and immediately frozen in liquid N₂ to avoid unwanted enzymatic reactions and stored at $-80^{\circ}C$ before the freeze-drying process. Lyophilized samples were powdered in liquid nitrogen and stored at $-20^{\circ}C$.

Extraction procedure for metabolomic analysis. An aliquot (50 mg) of freeze-dried and powdered plant material was transferred to a 2 mL microtube. 1.5 mL of phosphate buffer (Fluka Chemika, Buchs, Switzerland; 90 mM; pH 6.0) in D₂O (Cambridge Isotope Laboratories, Andover, MA, USA) - containing 0.1% w/w trimethylsilylpropionic-2,2,3,3-d₄ acid sodium salt (TMSP, Sigma-Aldrich, St. Louis, MO, USA)- and CD₃OD (Sigma-Aldrich, St. Louis, MO, USA) (1:1) were added to the samples. The mixture was vortexed at room temperature for 1 min, ultrasonicated (Elma Transsonic Digital, Hohentwiel, Germany) for 40 min, and centrifuged (Beckman Allegra™ 64 R, F2402H rotor; Beckman Coulter, Fullerton, CA, USA) at 13,000 rpm for 10 min. A volume of 0.65 mL was transferred to a 5-mm NMR tube and analyzed by NMR⁵⁶.

NMR experiments. NMR spectra were recorded at 25 °C on a 300.03 MHz for ¹H and 75.45 MHz for ¹³C on a Varian Mercury Plus 300 Fourier transform NMR. CD₃OD was used as the internal lock. Each ¹H NMR spectrum consisted of 256 scans with the following parameters: 0.16 Hz/point, acquisition time (AQ) = 1.0 s, relaxation delay (RD) = 1.5 s, 90° pulse width (PW) = 13.8 μs. A presaturation sequence was used to suppress the residual H₂O signal. FIDs were Fourier transformed with LB = 0.3 Hz. The spectra were manually phased and baseline-corrected and calibrated to TMSP at 0.0 ppm.

High-resolution experiments were performed, in particular 2D-NMR spectra were used for extracts *A. boeoticus* and *T. esculenta* extracts to identify the putative cytotoxic metabolites. Standard pulse sequences and phase cycling from Varian library were used for ¹H, DQF-COSY, COSY, TOCSY, HSQC, H2BC, HSQC TOCSY, HMBC and CIGAR-HMBC experiments. COSY spectra were acquired with a 1.0 s relaxation delay and 2514 Hz spectral width in both dimensions. The window function for COSY spectra was sine-bell (SSB = 0). Proton-detected heteronuclear correlations were measured. Heteronuclear single-quantum coherence (HSQC) experiments (optimized for ¹J_(H,C) = 140 Hz) were performed in the phase sensitive mode with field gradient. The spectral width was 12000 Hz in f1 (¹³C) and 3000 Hz in f2 (¹H) and 1.0 s of relaxation delay; the matrix of 1 k × 1 k data points was zero-filled to give a final matrix of 2 k × 2 k points. Heteronuclear 2 bond correlation (H2BC) spectra were obtained with T = 30.0 ms, and a relaxation delay of 1.0 s; the third-order low-pass filter was set for 130 < ¹J_(C,H) < 165 Hz. Heteronuclear multiple bond coherence (HMBC) experiment (optimized for ⁿJ_(H,C) = 8 Hz) was performed in the absolute value mode with field gradient. Typically, ¹H-¹³C gHMBC were acquired with

spectral width of 18116 Hz in f1 (^{13}C) and 3140 Hz in f2 (^1H) and 1.0 s of relaxation delay; the matrix of $1\text{ k} \times 1\text{ k}$ data points was zero-filled to give a final matrix of $4\text{ k} \times 4\text{ k}$ points; Qsine (SSB = 2.0). Constant time inverse-detection gradient accordion rescaled heteronuclear multiple bond correlation spectroscopy (CIGAR-HMBC) spectra ($8 > {}^n\text{J}_{(\text{H,C})} > 5$) were acquired with the same spectral width used for HMBC. Heteronuclear single-quantum coherence - total correlation spectroscopy (HSQC-TOCSY) experiments were optimized for ${}^n\text{J}_{(\text{H,C})} = 8\text{ Hz}$, with a mixing time of 90 ms. Free induction decays (FIDs) were Fourier transformed and the resulting spectra were manually phased and baseline-corrected and calibrated to TMS at 0.0 ppm, using ^1H NMR processor (ACDLABS 12.0, Toronto, Canada). For accurate mass measurements the purified compounds were analyzed by electrospray hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Q-TOF) fitted with a Z-spray electrospray ion source (Waters S.p.A.). All analyses were performed in both positive and negative ion mode. The capillary source voltage and the cone voltage were set at 3500 V and 35 V, respectively. The source temperature was kept at 80 °C and nitrogen was used as a drying gas (flow rate about 50 L/h). The time-of-flight analyzer of the mass spectrometer was externally calibrated with GFP from m/z 50 to 1600. Accurate mass data were collected by directly infusing samples (1.5 pmol/ μL in $\text{CH}_3\text{CN}/\text{H}_2\text{O}$, 1:1) into the system at a flow rate of 15 $\mu\text{L}/\text{min}$. The acquisition and processing of data were performed with the MassLynx 4.1 software (Waters S.p.A., Manchester, UK).

Isolation of active metabolites. *Extraction, isolation and identification of compound A from Astragalus boeoticus.* Dried leaves (24.0 g) of *A. boeoticus* were powdered and underwent to three cycles of an ultrasound assisted extraction with a $\text{MeOH}/\text{H}_2\text{O}$ (1:1) solution (720 mL), obtaining a crude extract (7.1 g). This was dissolved in H_2O and separated by liquid-liquid extraction using EtOAc as an extracting solvent. The EtOAc fraction (1 g) was chromatographed by SiO_2 CC eluting with a solution with an increasing degree of polarity (CHCl_3 , $\text{Me}_2\text{CO}/\text{CHCl}_3$, $\text{MeOH}/\text{CHCl}_3$). 21 fractions have been collected, number 12 was fractionated by Flash-CC eluting with $\text{MeOH}/\text{CHCl}_3$ (3:100) to obtain compound A (28 mg). Compound A (6-O-acetyl-3-O- β -D-xylopiranosylcycloastragenol). $[\alpha]_{\text{D}}^{25} = +7.65$ ($c = 5, 1 \times 10^{-3}$, $\text{MeOH}/\text{H}_2\text{O}$, 2:1). ^{13}C -NMR (CD_3OD): see Table S1; ^1H -NMR (CD_3OD): see Table S1. Q-TOF HRMS (ESI^+): m/z 703.3877 $[\text{M} + \text{K}]^+$ (calcd. 703.3818 for $\text{C}_{37}\text{H}_{60}\text{O}_{10}\text{K}^+$), 687.4080 $[\text{M} + \text{Na}]^+$ (calcd. 687.4079 for $\text{C}_{37}\text{H}_{60}\text{O}_{10}\text{Na}^+$), 627.3801 $[\text{M}-\text{AcOH} + \text{Na}]^+$ (calcd. 627.3873 for $\text{C}_{35}\text{H}_{56}\text{O}_8\text{Na}^+$), 537.3399 $[\text{M}-\text{C}_5\text{H}_{10}\text{O}_5 + \text{Na}]^+$ (calcd. 537.3556 $\text{C}_{32}\text{H}_{50}\text{O}_5\text{Na}^+$), 477.3164 $[\text{M}-\text{AcOH}-\text{C}_5\text{H}_{10}\text{O}_5 + \text{Na}]^+$ (calcd. 477.3345 $\text{C}_{30}\text{H}_{46}\text{O}_3\text{Na}^+$).

Extraction, isolation and identification of compound B from Trigonella esculenta. Dried leaves (10.0 g) of *T. esculenta* Willd were powdered and underwent to three cycles of an ultrasound assisted extraction with a $\text{MeOH}/\text{H}_2\text{O}$ (1:1) solution (300 ml) obtaining a crude extract (5.0 g), which was dissolved in H_2O and separated by liquid-liquid extraction using EtOAc as an extracting solvent. The aqueous fraction was chromatographed on Amberlite XAD-4 and XAD-7 with water and then with methanol. The alcoholic eluate from both XAD-4 and XAD-7 (1.2 g) was chromatographed by Sephadex LH-20 CC eluting with hydroalcoholic solution ($\text{MeOH}/\text{H}_2\text{O}$) with an increasing percentage of methanol to collect 8 fractions (20 mL each). Fractions 2 and 3 have been purified by Flash CC eluting with the lower phase of $\text{CHCl}_3/\text{MeOH}/\text{H}_2\text{O}$ (13:7:3) solution to give another fraction (74.0 mg), which in turn was chromatographed through HYDRO-RP 80 HPLC eluting with $\text{H}_2\text{O}/\text{MeOH}/\text{MeCN}$ (2:2:1) providing pure compound B (52.0). Compound B ([25 R]-furost-5-ene-3 β ,22 α ,26-triol 3-O- α -L-rhamnopyranosyl-(1 \rightarrow 4)- α -L-rhamnopyranosyl-(1 \rightarrow 4)-[α -L-rhamnopyranosyl-(1 \rightarrow 2)]- β -D-glucopyranosyl 26-O- β -D-glucopyranoside. $[\alpha]_{\text{D}}^{25} = +5.5$ ($c = 5, 1 \times 10^{-3}$, $\text{MeOH}/\text{H}_2\text{O}$, 2:1). ^{13}C -NMR (CD_3OD): S2; ^1H -NMR (CD_3OD): see Table S1;

Q-TOF HRMS (ESI^+): m/z 1217.5895 $[\text{M} + \text{Na}]^+$ (calcd. 1217.5926 for $\text{C}_{57}\text{H}_{94}\text{O}_{26}\text{Na}^+$), m/z 1177.6051 $[\text{M}-\text{OH}]^+$ (calcd. 1177.6006 for $\text{C}_{57}\text{H}_{93}\text{O}_{25}^+$); TANDEM MS of 1177.6051: 1015.5361 $[\text{M}-\text{OH}-\text{Glc}]^+$ (calcd. 1015.5478 for $\text{C}_{51}\text{H}_{83}\text{O}_{20}$), 869.4803 $[\text{M}-\text{OH}-\text{Glc}-\text{Rha}]^+$ (calcd. 869.4899 for $\text{C}_{45}\text{H}_{73}\text{O}_{16}$), 723.4020 $[\text{M}-\text{OH}-\text{Glc}-2\text{Rha}]^+$ (calcd. 723.4319 for $\text{C}_{39}\text{H}_{63}\text{O}_{12}$), 577.3481 $[\text{M}-\text{OH}-\text{Glc}-3\text{Rha}]^+$ (calcd. 577.3740 for $\text{C}_{33}\text{H}_{53}\text{O}_7$), 415.3236 $[\text{M}-\text{OH}-2\text{Glc}-3\text{Rha}]^+$ (calcd. 415.3212 for $\text{C}_{27}\text{H}_{43}\text{O}_3$).

Cell Lines. The human HCT-116, HT-29, Caco-2 colorectal cancer cell lines were obtained from the American Type Culture Collection (ATCC) (Manassas, VA). HCT-116, HT-29 cancer cells were cultured in RPMI 1640 medium (Lonza, Cologne, Germany) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 50 U/ml penicillin and 100 $\mu\text{g}/\text{ml}$ streptomycin (Lonza, Cologne, Germany). Caco-2 cell line was cultured in DMEM medium (Lonza, Cologne, Germany) supplemented with 10% fetal bovine serum, 2mM L-glutamine, 1% non-essential amino acid, 50 U/ml penicillin and 100 $\mu\text{g}/\text{ml}$ streptomycin (Lonza, Cologne, Germany).

Proliferation Assay. Cell proliferation assay was measured with 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. Briefly, cells in logarithmic growth phase were plated in 96-well plates and incubated for 24 hours before exposure to increasing doses of plant extracts (10, 50, 100, 150, 200 and 250 $\mu\text{g}/\text{ml}$), enriched in secondary metabolites by a Sep-pak C18 cartridge step. 48 hours after treatment, 50 μl of 1 mg/ml (MTT) were mixed with 200 μl of medium and added to the well. 1 hour after incubation at 37 °C, the medium was removed and the purple formazan crystals produced in the viable cells were solubilized in 100 μl of dimethyl sulfoxide and quantitated by measurement of absorbance at 570 nm with a plate reader. Results were reported as mean \pm sd of % of cell growth respect to the control from six replicate analyses. The control was represented by 0.25% DMSO treatment corresponding to the higher amount of DMSO used for the tests.

Multivariate data analysis. The data matrix of cell growth percentage from control of the plant extracts treatments has been analyzed by multivariate analysis based on XLSTAT 2013 (Addinsoft, New York, NY, USA for

Microsoft Office Excel 2010). Hierarchical cluster analysis (HCA), based on Euclidean distance as dissimilarity index and complete linkage as the agglomeration method, has been carried out to evaluate statistical differences between extracts. Furthermore, Principal component analysis (PCA) was also performed as ordination methods based on XLSTAT 2013.

References

- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod.* **75**(3), 311–35 (2012).
- Cragg, G. M., Newman, D. J. Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta.* **1830**(6) 3670–95 (2013).
- Mann, J. Natural products in cancer chemotherapy: past, present and future. *Nat Rev Cancer.* **2**(2), 143–8 (2002).
- Paterson, I. & Anderson, E. A. Chemistry. *The renaissance of natural products as drug candidates.* *Science* **310**(5747), 451–3 (2005).
- Wolfender, J.-L., Marti, G. & Ferreira Queiroz, E. Advances in Techniques for Profiling Crude Extracts and for the Rapid Identification of Natural Products: Dereplication, Quality Control and Metabolomics. *Current Organic Chemistry* **14**(16), 1808–32 (2010).
- Michel, T., Halabalaki, M. & Skaltsounis, A. L. New concepts, experimental approaches, and dereplication strategies for the discovery of novel phytoestrogens from natural sources. *Planta Med.* **79**(7), 514–32 (2013).
- Harvey, A. L. Natural products in drug discovery. *Drug Discov Today.* **13**(19–20), 894–901 (2008).
- Verpoorte, R., Choi, Y. H. & Kim, H. K. Ethnopharmacology and systems biology: a perfect holistic match. *J Ethnopharmacol.* **100**(1–2), 53–6 (2005).
- Nakabayashi, R. & Saito, K. Metabolomics for unknown plant metabolites. *Anal Bioanal Chem.* **405**(15), 5005–11 (2013).
- Lawaetz, A. J., Schmidt, B., Staerk, D., Jaroszewski, J. W. & Bro, R. Application of rotated PCA models to facilitate interpretation of metabolite profiles: commercial preparations of St. John's Wort. *Planta Med.* **75**(3), 271–9 (2009).
- Yuliana, N. D., Khatib, A., Choi, Y. H. & Verpoorte, R. Metabolomics for bioactivity assessment of natural products. *Phytother Res.* **25**(2), 157–69 (2011).
- Malvezzi, M. *et al.* European cancer mortality predictions for the year 2016 with focus on leukaemias. *Ann Oncol.* **27**(4), 725–31 (2016).
- Van Cutsem, E. *et al.* Cetuximab and Chemotherapy as Initial Treatment for Metastatic Colorectal Cancer. *New Engl J Med.* **360**(14), 1408–17 (2009).
- Misale, S. *et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* **486**(7404), 532–6 (2012).
- Troiani, T. *et al.* Increased TGF- α as a mechanism of acquired resistance to the anti-EGFR inhibitor cetuximab through EGFR-MET interaction and activation of MET signaling in colon cancer cells. *Clin Cancer Res.* **19**(24), 6751–65 (2013).
- Wink, M. Evolution of secondary metabolites in legumes (Fabaceae). *South African Journal of Botany* **89**, 164–75 (2013).
- Dixon, R. A. & Sumner, L. W. Legume natural products: understanding and manipulating complex pathways for human and animal health. *Plant Physiol.* **131**(3), 878–85 (2003).
- Toyomura, K. & Kono, S. Soybeans, Soy Foods, Isoflavones and Risk of Colorectal Cancer: a Review of Experimental and Epidemiological Data. *Asian Pac J Cancer Prev.* **3**(2), 125–32 (2002).
- Auyeung, K. K., Law, P. C. & Ko, J. K. Novel anti-angiogenic effects of formononetin in human colon cancer cells and tumor xenograft. *Oncol Rep.* **28**(6), 2188–94 (2012).
- Kang, H. W. *et al.* Deguelin, an Akt inhibitor, down-regulates NF- κ B signaling and induces apoptosis in colon cancer cells and inhibits tumor growth in mice. *Dig Dis Sci.* **57**(11), 2873–82 (2012).
- Lee, C. K. *et al.* Isoliquiritigenin inhibits tumor growth and protects the kidney and liver against chemotherapy-induced toxicity in a mouse xenograft model of colon carcinoma. *J Pharmacol Sci.* **106**(3), 444–51 (2008).
- Ellington, A. A., Berhow, M. A. & Singletary, K. W. Inhibition of Akt signaling and enhanced ERK1/2 activity are involved in induction of macroautophagy by triterpenoid B-group soyasaponins in colon cancer cells. *Carcinogenesis.* **27**(2), 298–306 (2006).
- D'Abrosca, B. *et al.* Plant bioassay to assess the effects of allelochemicals on the metabolome of the target species *Aegilops geniculata* by an NMR-based approach. *Phytochemistry.* **93**, 27–40 (2013).
- Esposito, A. *et al.* Potential allelopathic interference of *Melilotus neapolitana* metabolites on three coexisting species of Mediterranean herbaceous plant community. *Journal of Plant Interactions* **3**, 199–210 (2008).
- Kitagawa, I., Wang, H., Saito, M., Takagi, A. & Yoshikawa, M. Saponin and sapogenol. XXXV. Chemical constituents of astragali radix, the root of *Astragalus membranaceus* Bunge. (2). Astragalosides I, II and IV, acetylastragaloside I and isoastragalosides I and II. *Chemical & Pharmaceutical Bulletin.* **31**(2), 698–708 (1983).
- Polat, E., Bedir, E., Perrone, A., Piacente, S. & Alankus-Caliskan, O. Triterpenoid saponins from *Astragalus wiedemannianus* Fischer. *Phytochemistry.* **71**(5–6), 658–62 (2010).
- Murakami, T., Kishi, A., Matsuda, H. & Yoshikawa, M. Medicinal Foodstuffs. XVII. Fenugreek Seed. (3). Structures of New Furostanol-Type Steroid Saponins, Trigoneosides Xa, Xb, Xib, XIa, XIb, and XIIIa, from the Seeds of Egyptian *Trigonella foenum-graecum* L. *Chemical & Pharmaceutical Bulletin.* **48**(7), 994–1000 (2000).
- Wang, Z. B. *et al.* Triterpenoids and Flavonoids from the Leaves of *Astragalus membranaceus* and Their Inhibitory Effects on Nitric Oxide Production. *Chem Biodivers.* **12**(10), 1575–84 (2015).
- Munday, S. C., Wilkins, A. L., Miles, C. O. & Holland, P. T. Isolation and structure elucidation of dichotomin, a furostanol saponin implicated in hepatogenous photosensitization of sheep grazing *Panicum dichotomiflorum*. *Journal of Agricultural and Food Chemistry.* **41**(2), 267–71 (1993).
- Agrawal, P. K. Assigning stereodiversity of the 27-Me group of furostane-type steroidal saponins via NMR chemical shifts. *Steroids.* **70**(10), 715–24 (2005).
- Balunas, M. J. & Kinghorn, A. D. Drug discovery from medicinal plants. *Life Sci.* **78**(5), 431–41 (2005).
- Kinghorn, A. *et al.* Novel Strategies for the Discovery of Plant-Derived Anticancer Agents. *Pharmaceutical Biology.* **41**, 53–67 (2008).
- Verpoorte, R., Choi, Y. H., Mustafa, N. R. & Kim, H. K. Metabolomics: back to basics. *Phytochem Rev.* **7**(3), 525–37 (2008).
- Khoogar, R., Kim, B. C., Morris, J. & Wargovich, M. J. Chemoprevention in gastrointestinal physiology and disease. Targeting the progression of cancer with natural products: a focus on gastrointestinal cancer. *Am J Physiol Gastrointest Liver Physiol.* **310**(9), G629–44 (2016).
- Hanbury, C. D., White, C. L., Mullanc, B. P. & Siddiquea, K. H. M. A review of the potential of *Lathyrus sativus* L. and *L. cicera* L. grain for use as animal feed. *Animal Feed Science and Technology* **87**, 1–27 (2000).
- Nithiyantham, S., Selvakumar, S. & Siddhuraju, P. Total phenolic content and antioxidant activity of two different solvent extracts from raw and processed legumes. *Cicer arietinum* L. and *Pisum sativum* L. *Journal of Food Composition and Analysis* **27**, 52–60 (2012).
- Wang, N. & Daun, J. K. Effect of variety and crude protein content on nutrients and certain antinutrients in field peas. *J Sci Food Agric* **84**, 1021–1029 (2004).
- Sammour, H. R., Mustafa, A. E., Badr, S. & Tahr, W. Genetic variability of some quality traits in *Lathyrus* spp. Germoplasm. *Acta agric Slovenica* **90**, 33–43 (2007).

39. Li, X. *et al.* A review of recent research progress on the astragalus genus. *Molecules*. **19**(11), 18850–80 (2014).
40. Lee, D. Y. *et al.* Anti-inflammatory cycloartane-type saponins of *Astragalus membranaceus*. *Molecules*. **18**(4), 3725–32 (2013).
41. Bagherwal, P. Phytosaponin Adjuvants: A Better Option for Vaccines. *International Journal of PharmTech Research*. **3**, 1837–42 (2011).
42. Yesilada, E., Bedir, E., Calis, I., Takaishi, Y. & Ohmoto, Y. Effects of triterpene saponins from *Astragalus* species on *in vitro* cytokine release. *J Ethnopharmacol*. **96**(1–2), 71–7 (2005).
43. Auyeung, K. K., Woo, P. K., Law, P. C. & Ko, J. K. *Astragalus* saponins modulate cell invasiveness and angiogenesis in human gastric adenocarcinoma cells. *J Ethnopharmacol*. **141**(2), 635–41 (2012).
44. Varshney, I. P., Sood, A. R., Srivastava, H. C. & Harshe, S. N. Isolation of ethyl galactoside from *Trigonella esculenta* Willd seeds. *Planta Med*. **26**(1), 26–32 (1974).
45. Mahna, S., Raisinghani, G. & Jain, S. Diosgenin production in induced mutants of *Trigonella esculenta* Willd. *Fitoterapia*. **65**(6), 515–6 (1994).
46. Yadav, U. C. & Baquer, N. Z. Pharmacological effects of *Trigonella foenum-graecum* L. in health and disease. *Pharm Biol*. **52**(2), 243–54 (2014).
47. Mehrafarin, A. *et al.* Bioengineering of Important Secondary Metabolites and Metabolic Pathways in Fenugreek (*Trigonella foenum-graecum* L.). *Journal of Medicinal Plants*. **3**(35), 1–18 (2010).
48. Basch, E., Ulbricht, C., Kuo, G., Szapary, P. & Smith, M. Therapeutic applications of fenugreek. *Altern Med Rev*. **8**(1), 20–7 (2003).
49. Shabbeer, S. *et al.* Fenugreek: A naturally occurring edible spice as an anticancer agent. *Cancer Biology & Therapy*. **8**(3), 272–8 (2014).
50. Prabhu, A. & Krishnamoorthy, M. Anticancer activity of *Trigonella foenum-graecum* on Ehrlich Ascites carcinoma in *Mus musculus* system. *Journal of Pharmacy Research*. **3**(6), 1181–3 (2010).
51. Ionkova, I., Momekov, G. & Proksch, P. Effects of cycloartane saponins from hairy roots of *Astragalus membranaceus* Bge., on human tumor cell targets. *Fitoterapia*. **81**(5), 447–51 (2010).
52. Huang, C. *et al.* Reversal of P-glycoprotein-mediated multidrug resistance of human hepatic cancer cells by Astragaloside II. *J Pharm Pharmacol*. **64**(12), 1741–50 (2012).
53. Van Cutsem, E., Cervantes, A., Nordlinger, B., Arnold, D. & Group, E. G. W. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. **25**, 1–9 (2014).
54. Van Cutsem, E. *et al.* Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. *J Clin Oncol*. **29**(15), 2011–9 (2011).
55. Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature*. **483**(7387), 100–U46 (2012).
56. Kim, H. K., Choi, Y. H. & Verpoorte, R. NMR-based metabolomic analysis of plants. *Nat. Protoc*. **5**, 536–549 (2010).

Author Contributions

V.G. and M.S. designed the experiments, discussed and interpreted the results. A.E. selected and identified the plant species. V.G., M.S., V.B., and M.P. carried out experiments. A.C. and R.R. carried out the M.S. analyses. V.G. and M.S. wrote the manuscript. A.E., B.D.A., A.R., F.C. analysed the experimental data. A.E., B.D.A. interpreted the results. T.T., N.P., A.F. designed the experiments, discussed and interpreted the results and reviewed the manuscript. All authors revised the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-23704-9>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018