Design of Novel Drug-like Molecules using Informatics Rich Secondary Metabolites Analysis of Indian Medicinal and Aromatic Plants

**Supplementary file S2.1:** Computational Protocol for Identification of ‘specific’ metabolites and building ‘virtual library’ with predicted bioactivity

**Case study:** Query list of Indian medicinal plants are the following:

1. Indian medicinal plants falling under none of the category of four systems of medicine- Count: 5)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DoMINE (104 plants)** | **Ayurveda** | **Homeopathy** | **Siddha** | **UNANI** | **Total** |
| *Rosmarianus officinalis* | 0 | 0 | 0 | 0 | *0* |
| *Pelargonium graveolens* | 0 | 0 | 0 | 0 | *0* |
| *Tagetes minuta* | 0 | 0 | 0 | 0 | *0* |
| *Cymbopogon winterianus* | 0 | 0 | 0 | 0 | *0* |
| *Dracocephalum heterophyllum* | 0 | 0 | 0 | 0 | *0* |

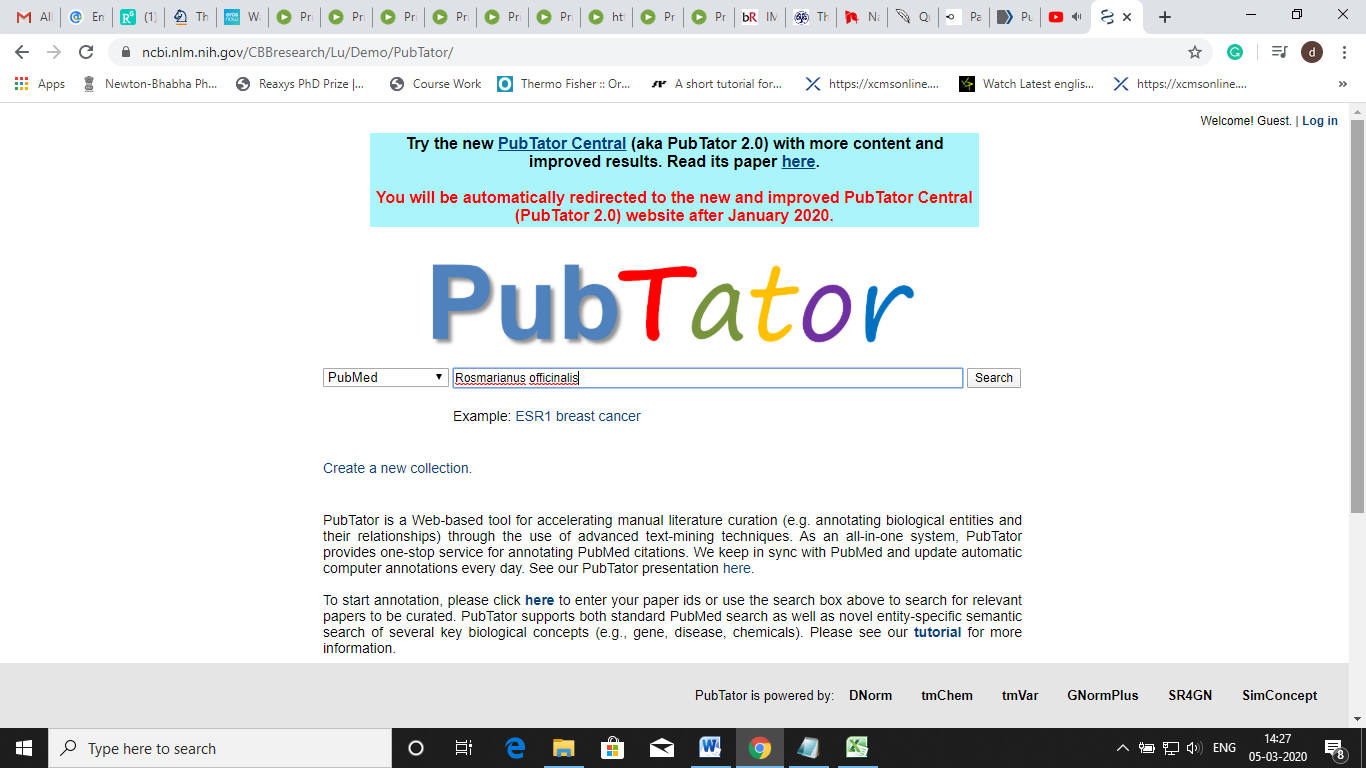
1. Indian medicinal plants falling under some of the four systems of medicine-Count: 3)

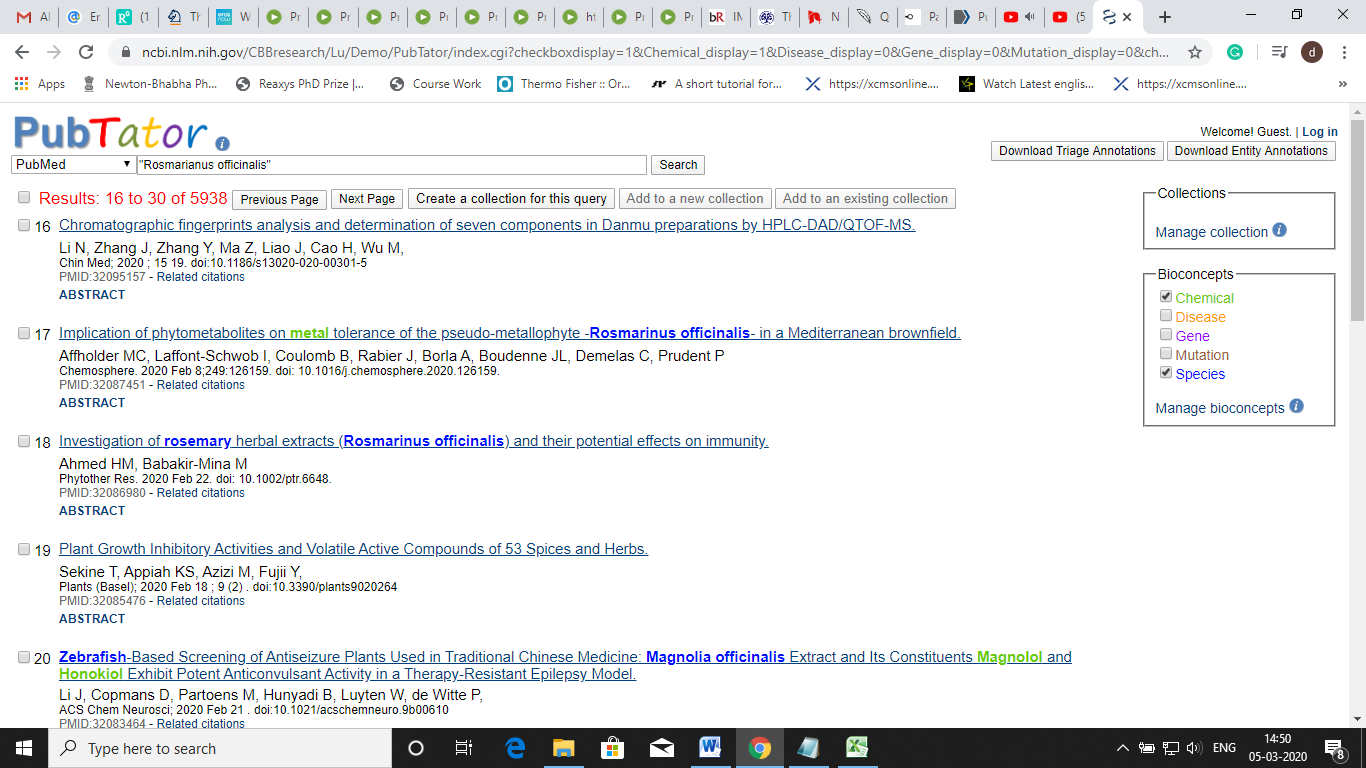
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DoMINE (104 plants)** | **Ayurveda** | **Homeopathy** | **Siddha** | **UNANI** | **Total** |
| *Pogostemon cablin* | 1 | 0 | 0 | 0 | 1 |
| *Artemisia maritima* | 1 | 1 | 0 | 0 | 2 |
| *Cymbopogon martini* | 1 | 0 | 1 | 0 | 2 |
| *Salvia sclarea* | 1 | 0 | 0 | 1 | 2 |

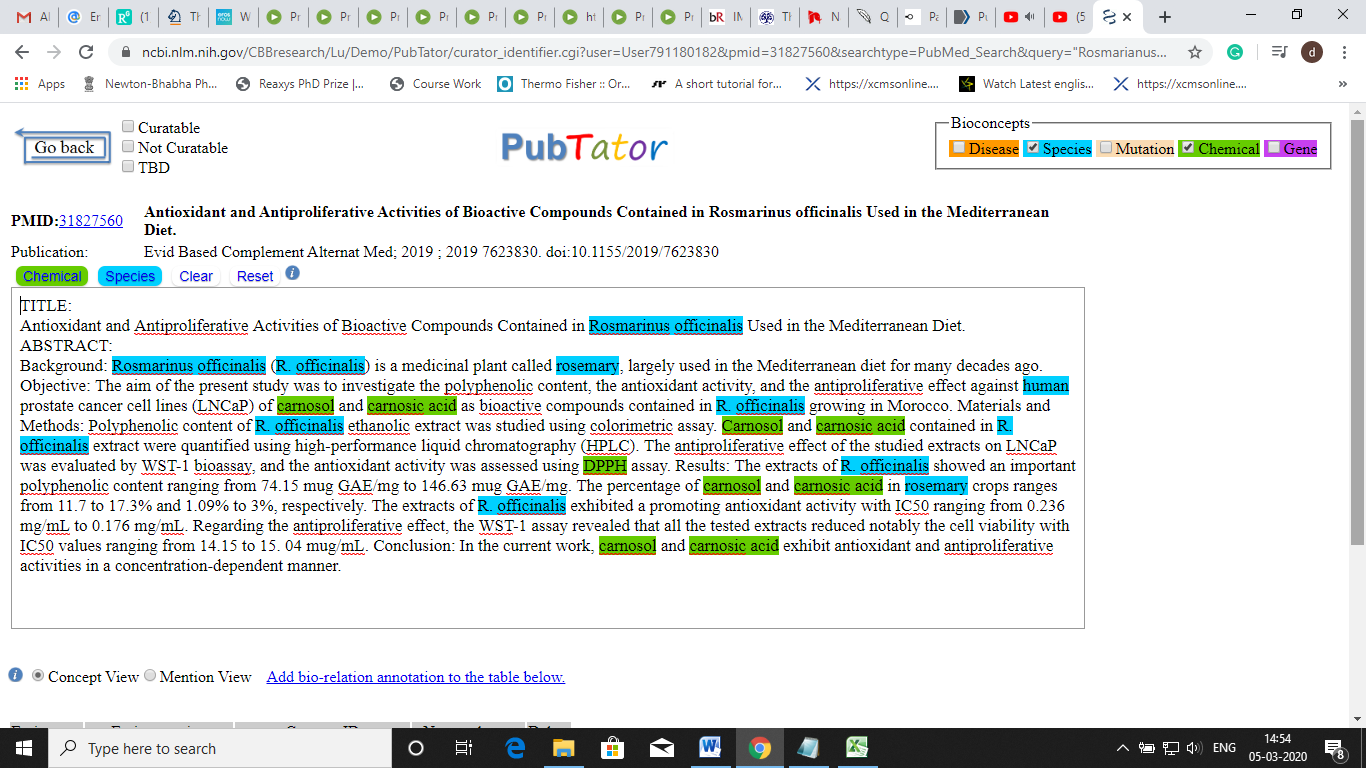
1. Indian medicinal plants falling under all four systems of medicnes. From them Select best 5 Indian medicinal plants- Count:5

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DoMINE (104 plants)** | **Ayurveda** | **Homeopathy** | **Siddha** | **UNANI** | **Total** |
| *Cannabis sativa* | 1 | 1 | 1 | 1 | 4 |
| *Linum usitatissimum* | 1 | 1 | 1 | 1 | 4 |
| *Croton tiglium* | 1 | 1 | 1 | 1 | 4 |
| *Ruta graveolens* | 1 | 1 | 1 | 1 | 4 |
| *Aconitum napellus* | 1 | 1 | 1 | 1 | 4 |

1. Textmining using pubTator (webtool):





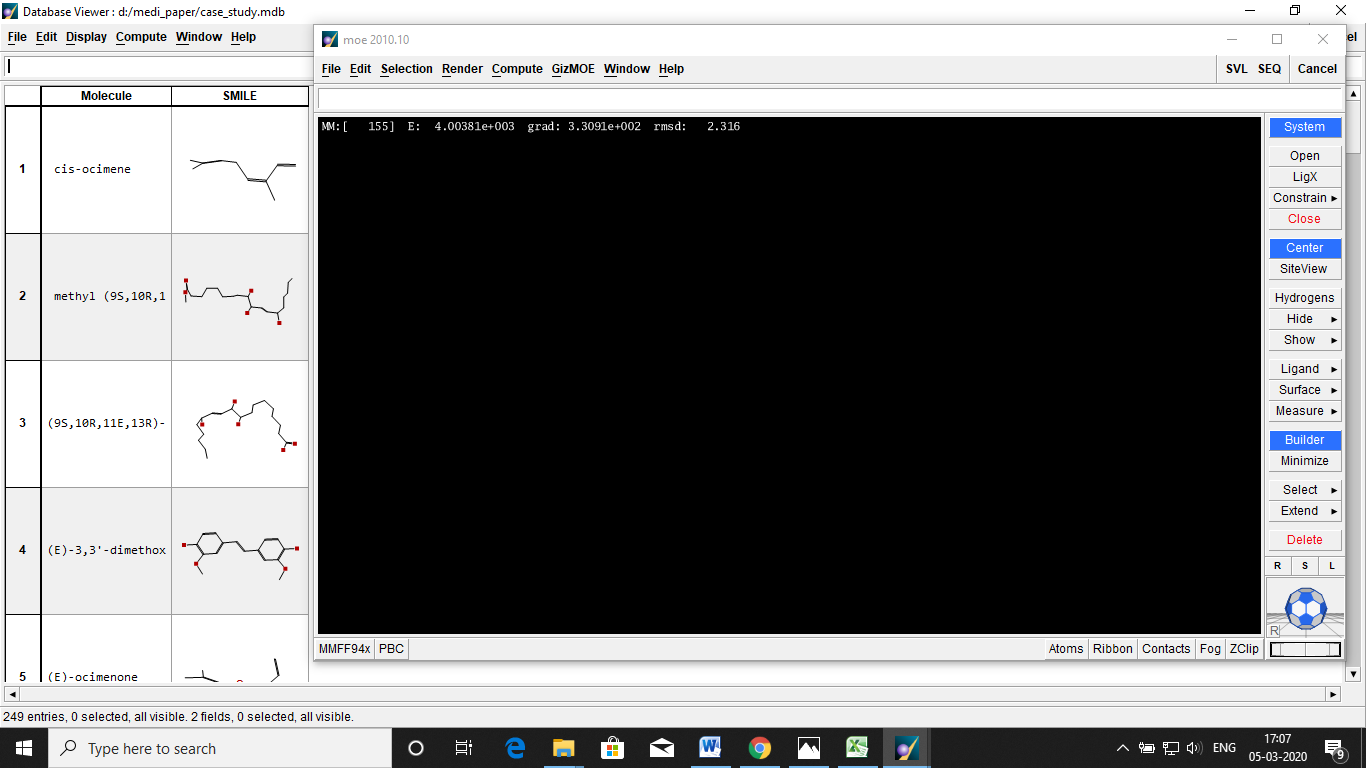


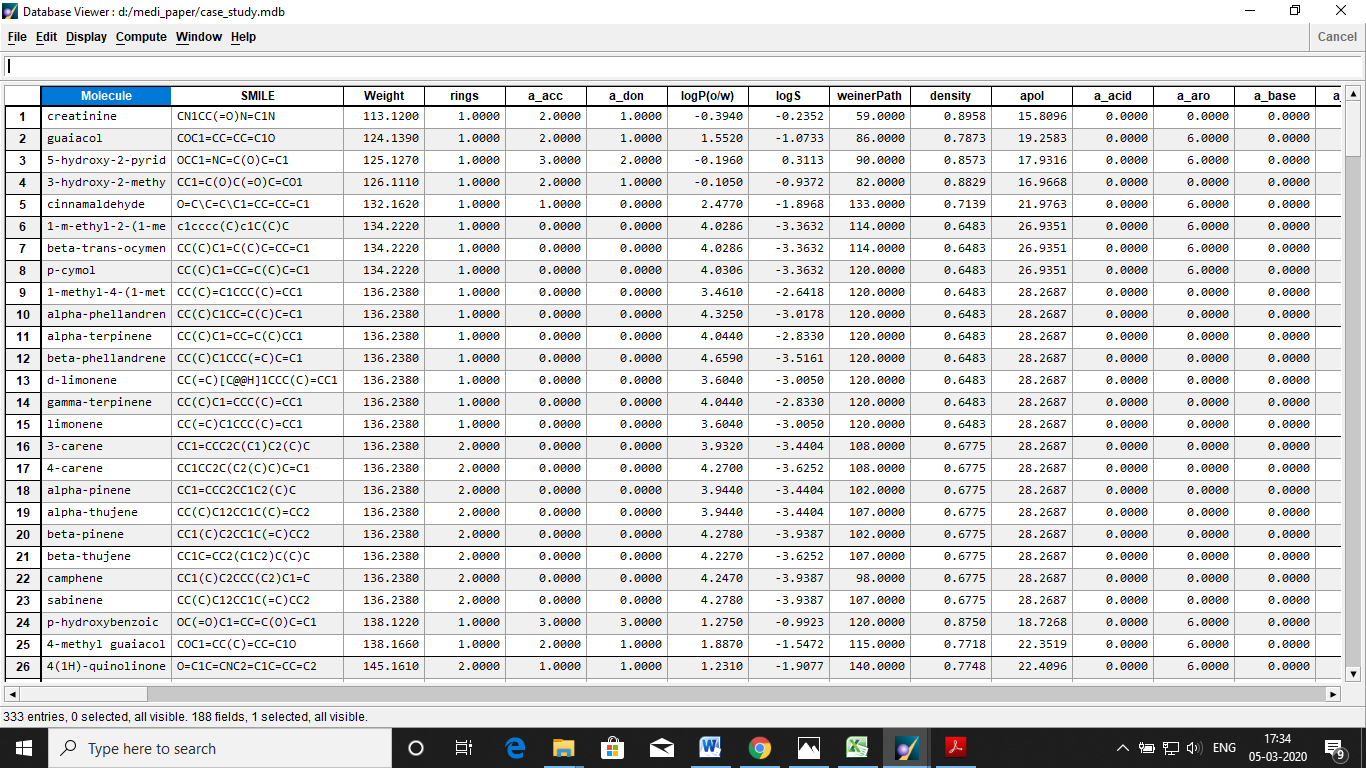
1. List of molecules (n=410) from the query list after downloading file in notepad and allighned in excel sheet.

*Rosmarianus officinalis:*

|  |  |  |  |
| --- | --- | --- | --- |
| PUBMED ID | Molecule | Class | ID |
| 31058853 | carnosic acid | Chemical | C018381 |
| 30994346 | abietane diterpene | Chemical |  |
| 30923446 | caffeic acids | Chemical |  |
| 30923446 | rosmarinic acid | Chemical | C041376 |
| 30923446 | apigenin | Chemical |  |
| 30923446 | luteolin | Chemical |  |
| 30881726 | 1,8-cineole | Chemical | C010087 |
| 30881726 | camphor | Chemical |  |
| 30486296 | chlorogenic acid | Chemical |  |
| 30486296 | gallic acid | Chemical | CHEBI:9505 |
| 30486296 | ferulic acid | Chemical |  |
| 30478939 | asiatic acid | Chemical | C017032 |
| 30478939 | vanillin | Chemical | C100058 |
| 30292288 | rosmanol | Chemical | C465581 |
| 30219443 | carnosolic acid | Chemical |  |
| 30219443 | trans-caffeic acid | Chemical | CHEBI:16433 |
| 30040945 | beta-myrcene | Chemical | C008574 |
| 30005652 | α-pinene | Chemical |  |
| 29915496 | camphene | Chemical |  |
| 29915496 | b-thujene | Chemical |  |
| 29915496 | a-thujene | Chemical |  |
| 29915496 | b-cubenene | Chemical |  |
| 29915496 | chrysanthenone | Chemical | C441667 |
| 28983254 | borneol | Chemical | C022871 |
| 28983254 | β-pinene | Chemical |  |
| 28983254 | verbenone | Chemical |  |
| 28895237 | feruloylnepitrin | Chemical |  |
| 28895237 | luteolin-3'-O-(2″-O-acetyl)-β-d-glucuronide | Chemical |  |
| 28895237 | luteolin-7-O-rutinoside | Chemical |  |
| 28802239 | ursolic acid | Chemical | C005466 |
| 28763707 | b-myrcene | Chemical |  |
| 28763707 | sabinene | Chemical |  |
| 28362473 | linoleic acid | Chemical | D019787 |
| 28106756 | officinoflavonosides A | Chemical |  |
| 28106756 | officinoflavonosides B | Chemical |  |
| 27823632 | sageone | Chemical |  |
| 30549611 | piperitenone oxide | Chemical |  |
| 30549611 | y-terpinene | Chemical |  |
| 25910439 | luteolin-7-O-glucuronide | Chemical | C456096 |
| 25631514 | terpinen-4-ol | Chemical |  |
| 25200369 | (1S,4S,5S)-5-exo-hydrocamphor 5-O-β-D-glucopyranoside | Chemical |  |
| 25200369 | officinoterpenosides A1 | Chemical |  |
| 25200369 | officinoterpenosides A2 | Chemical |  |
| 25200369 | isorosmanol | Chemical |  |
| 25200369 | epirosmanol | Chemical |  |
| 25200369 | micromeric acid | Chemical |  |
| 25200369 | glucosyl tormentate | Chemical |  |
| 25200369 | asteryunnanoside B | Chemical |  |
| 25200369 | niga-ichigoside F₁ | Chemical |  |
| 25200369 | oleanolic acid | Chemical |  |
| 25200369 | 7-methoxyrosmanol | Chemical | C510922 |
| 25200369 | officinoterpenosides B | Chemical |  |
| 25200369 | officinoterpenosides C | Chemical |  |
| 25200369 | officinoterpenosides d | Chemical |  |
| 24616107 | cis-verbenol | Chemical |  |
| 24616107 | myrcene | Chemical | C509595 |
| 24616107 | α-terpinene | Chemical |  |
| 24616107 | γ-terpinene | Chemical |  |
| 24616107 | α-terpineol | Chemical |  |
| 24043430 | (S)-(-)-cis-verbenol | Chemical |  |
| 24043430 | (S)-(-)-verbenone | Chemical |  |
| 23687854 | (+)-4-carene | Chemical |  |
| 23687854 | pinene | Chemical | CHEBI:17187 |
| 23687854 | isoborneol | Chemical | C022871 |
| 23687854 | eucalyptol | Chemical |  |
| 23687854 | limonene | Chemical |  |
| 23687854 | 1-methyl-4-(1-methylethylidene)-cyclohexene | Chemical |  |
| 23472478 | beta-phellandrene | Chemical | C058582 |
| 23472478 | bornyl acetate | Chemical |  |
| 24584866 | tannic acid | Chemical |  |
| 24584866 | cineol | Chemical |  |
| 22927089 | 7-O-methylrosmanol | Chemical |  |
| 22927089 | royleanonic acid | Chemical |  |
| 22424272 | rosmaridiphenol (11,12-dihydroxy-8,11,13-icetexatrien-1-one) | Chemical | C573384 |
| 21429498 | berbonone | Chemical |  |
| 21328358 | L-carvone | Chemical | CHEBI:15400 |
| 21151161 | carnasol | Chemical |  |
| 21138060 | beta-caryophyllene | Chemical | C024714 |
| 20397728 | 6''-O-(E)-feruloylhomoplantaginin | Chemical |  |
| 20397728 | 6''-O-(E)-feruloylnepitrin | Chemical |  |
| 20397728 | 6''-O-(E)-p-coumaroylnepitrin | Chemical |  |
| 20397728 | 6-methoxyluteolin 7-glucopyranoside | Chemical |  |
| 20397728 | luteolin 3'-O-(3''-O-acetyl)-beta-D-glucuronide | Chemical |  |
| 20397728 | 1-O-feruloyl-beta-D-glucopyranose | Chemical |  |
| 20397728 | 1-O-(4-hydroxybenzoyl)-beta-D-glucopyranose | Chemical |  |
| 20397728 | luteolin 3'-O-beta-D-glucuronide | Chemical |  |
| 20397728 | kaempferol | Chemical | C012843 |
| 20397728 | genkwanin | Chemical | D010118 |
| 20397728 | ladanein | Chemical | D019207 |
| 19184968 | borneol acetate | Chemical |  |
| 19184968 | 2-bornanone | Chemical | CHEBI:36773 |
| 18800806 | cis-4-glucosyloxycinnamic acid | Chemical |  |
| 18800806 | 3,4,5-trimethoxyphenylmethanol | Chemical |  |
| 15950250 | 7beta-methoxyabieta-8,13-diene-11,12-dione-(20,6beta)-olide (rosmaquinone A) | Chemical | C505171 |
| 15950250 | 7alpha-methoxyabieta-8,13-diene-11,12-dione-(20,6beta)-olide (rosmaquinone B) | Chemical | C505171 |
| 15561190 | 12-methoxy-trans-carnosic acid | Chemical |  |
| 15561190 | 12-methoxy-cis-carnosic acid | Chemical |  |
| 15291464 | isoscutellarein 7-O-glucoside | Chemical |  |
| 15291464 | hispidulin 7-O-glucoside | Chemical |  |
| 15291464 | eriocitrin | Chemical |  |
| 15291464 | hesperidin | Chemical |  |
| 15291464 | diosmin | Chemical |  |
| 7765765 | luteolin 3'-O-(4"-O-acetyl)-beta-D-glucuronide | Chemical |  |
| 7765765 | luteolin 3'-O-(3"-O-acetyl)-beta-D-glucuronide | Chemical |  |
| 5249317 | 5-hydroxy-7,4'-dimethoxyflavone | Chemical | C044998 |

1. Filtering of molecules in MOE for a unique list of molecules with an aromatic ring >= 6 and moleculer weight >=1000 MW. (n=330/410)





1. Scaffold and functional groups extraction from n=330 molecules using ChemScreener.

Extracted n=56 scaffolds and n= 32 functional groups.



**Table 1. 2-D structures of extracted scaffolds (n=15) from n=330 molecules**

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

1. Virtual library (n=510) generation from scaffolds (n=5) and functional groups (n=5)

**Table 2: 2-D structures of molecules and their scaffolds**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr. No.** | **Molecule Name** | **Molecule structure** | **Scaffold** |
|  | Beta-cedrene |  |  |
|  | licarin B |  |  |
|  | backuchiol |  |  |
|  | trans-cinnamic acid |  |  |
|  | gamma-fagarine |  |  |

1. Virtual library screening with T- toxicophore, P- Pharmacophore and Chemophore scores. Select molecules having more Pharmacophore and less toxicophores scores. Chemophore scores should also be less than Pharmacophore scores.

**Table 2: 2-D structures of the virtual library of novel molecules (n=10) with their TPC scores**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr. no.** | **Virtual library molecule** | **Toxicophore score** | **Pharmacophore score** | **Chemphore score** |
|  |  | 14 | 26 | 9 |
|  |  | 14 | 23 | 9 |
|  |  | 14 | 23 | 9 |
|  |  | 16 | 31 | 15 |
|  |  | 16 | 30 | 14 |
|  |  | 16 | 30 | 14 |
|  |  | 16 | 30 | 15 |
|  |  | 16 | 30 | 15 |
|  |  | 16 | 30 | 15 |
|  |  | 16 | 28 | 14 |