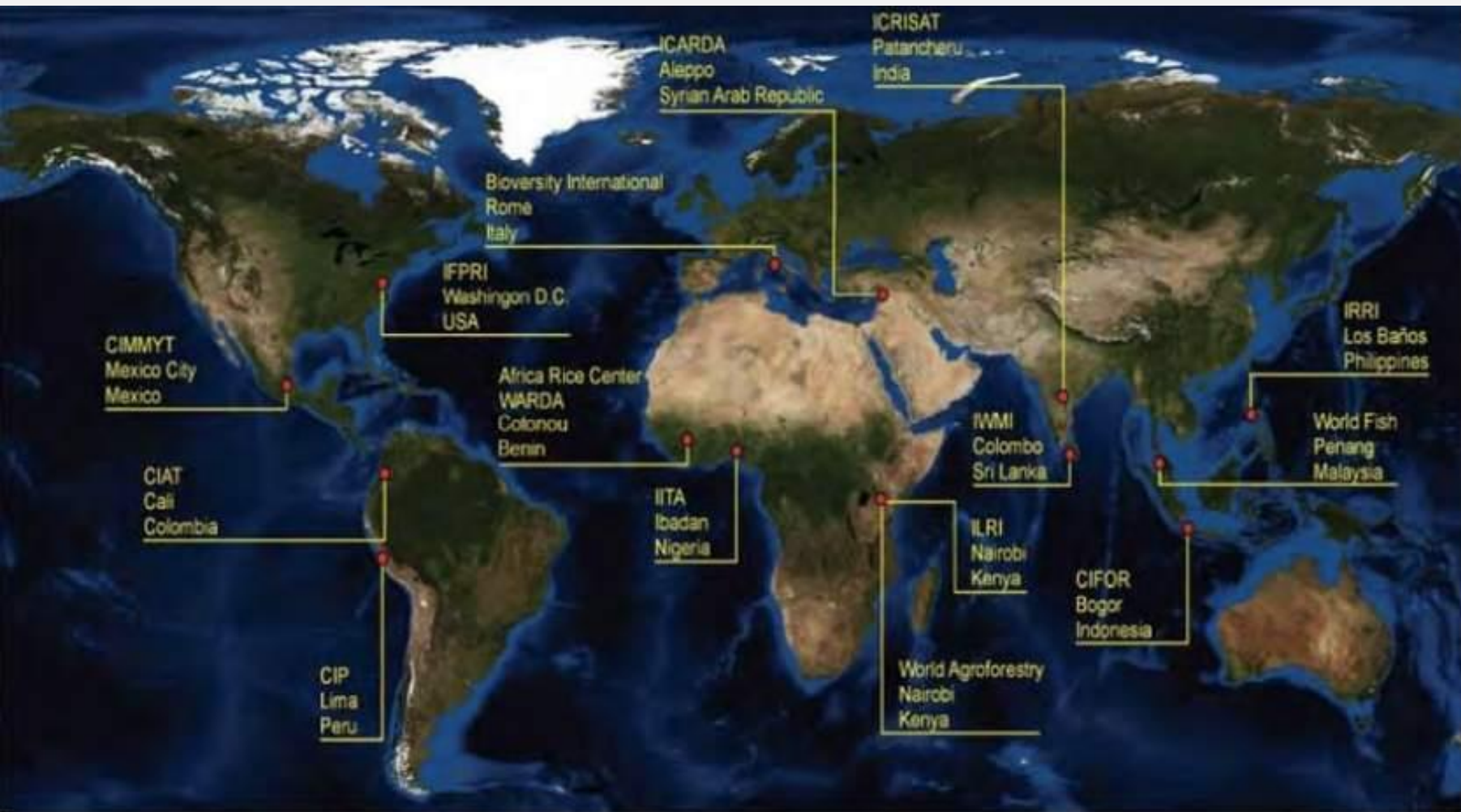


CGIAR's Big Data Platform



Medha Devare (m.devare@cgiar.org)
CAPSELLA Open Data Workshop, Chania | June 2, 2017

Who?



GOALS

Reduced poverty

Improved food and nutrition security for health

Improved natural resource systems and ecosystem services

15 Centers; >8000 scientists; fundamental to applied, multi-scale, multi-disciplinary research

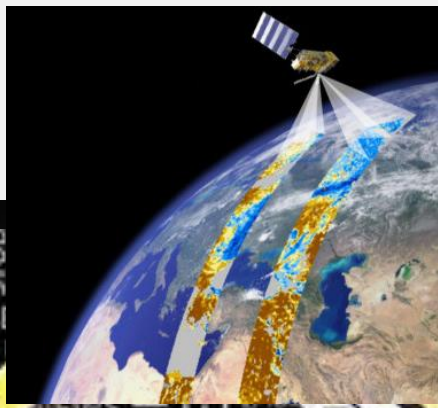
Why?

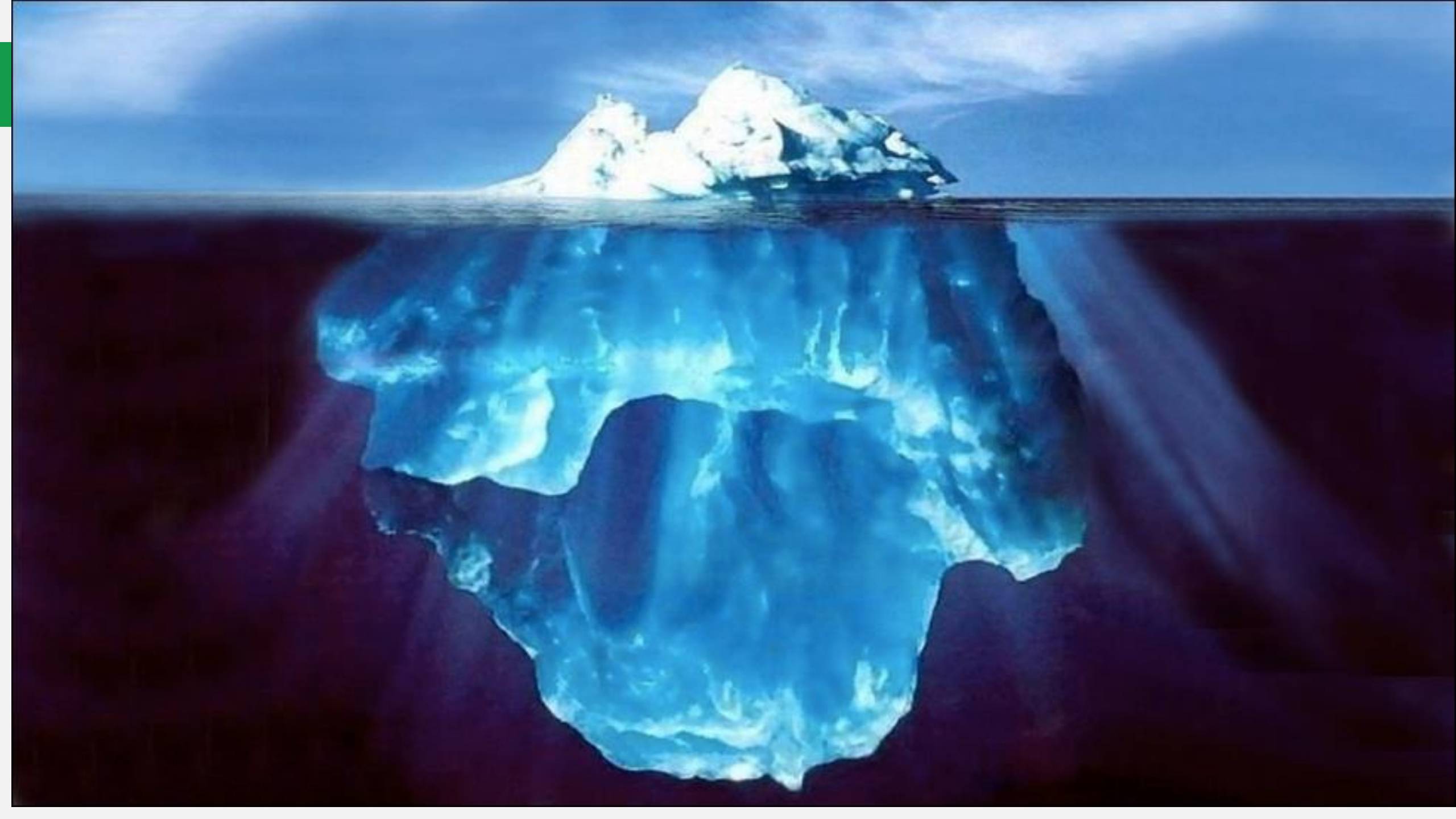
“The data and knowledge products generated by CGIAR arguably are assets of comparable social value to the content of the genebanks, which strongly suggests that **CGIAR has dramatically underinvested** in the curation and maintenance of these assets.

“The field is changing so fast that the only way to stay on the edge is to be invested and involved in these processes, which occur **outside** the CGIAR.

“...growing amount of frustration...on lack of data transparency and discoverability on breeding, remote sensing, soil agronomy, etc.”

Opportunit









Requires **FAIR datasets** → harmonization on **tools/platforms/standards**,
and... **incentives/culture**

Organize



Support data generation and management, practices, enable FAIR data

Convene



Bring together big data practitioners, private sector, academia, (I)NGOs++ to produce innovative solutions

Inspire



Employ analytics and ICTs to deliver info to farmers, monitor ag and food security, and inform policy



Issues: Culture, technical...business model...

- Address standardization at data collection – not at data storage stage (AMS)
- Support adoption of minimum metadata standard (CG Core)
- Establish and foster uptake of interoperability standards and protocols – through Communities of Practice where needed (controlled vocabularies, ontologies)
- Develop easy workflows, and data documentation, curation, QA/QC processes
- Ensure ease of use, value proposition for databases and tools
- Provide continuous support for researchers (advocacy, training, DMP help...)
- Incentivize data sharing (annual evals, “data sprints” with rewards, dataset citations, funding consequences...)



Organize

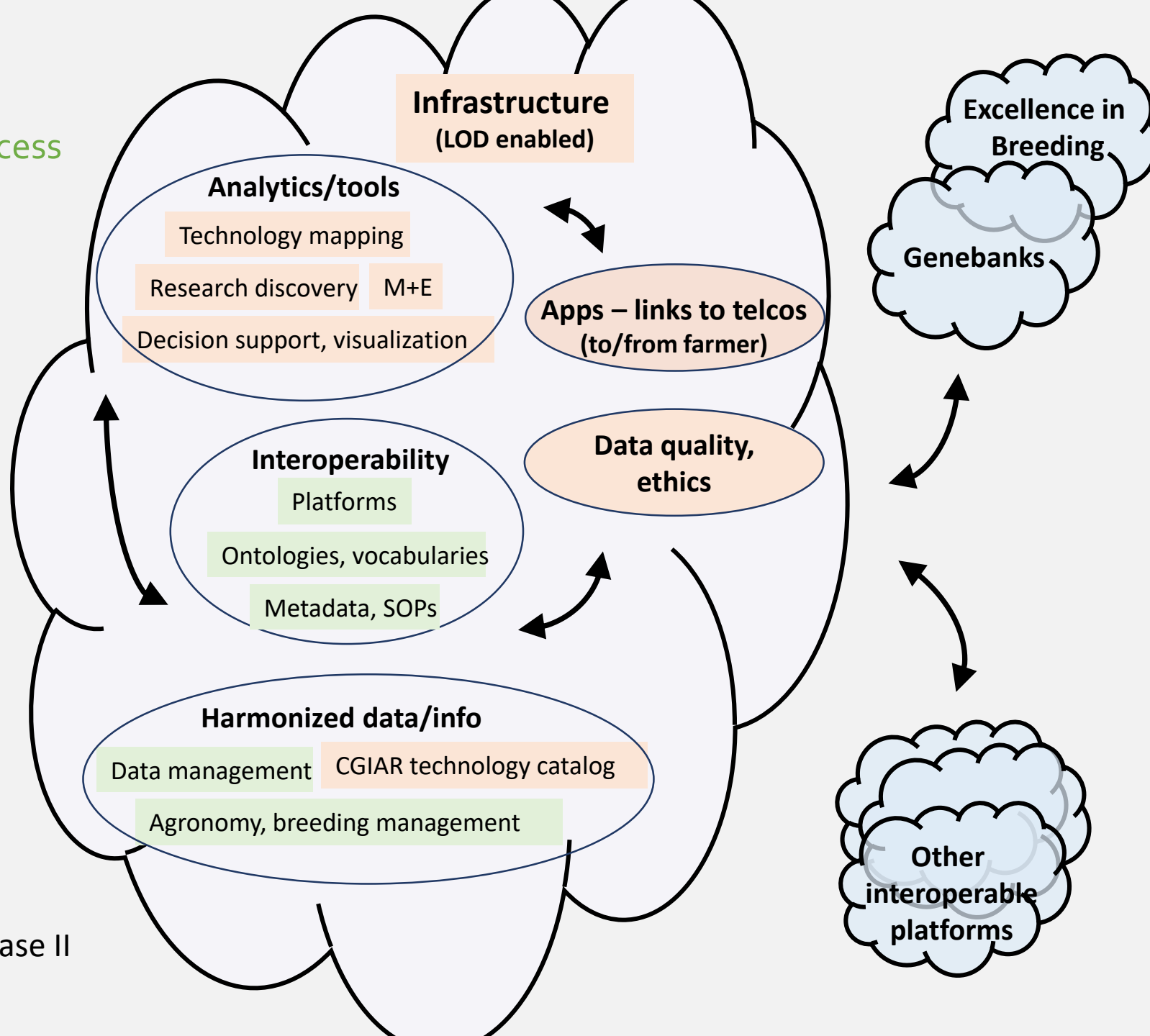
Support and improve data generation, management, access

Aligned donor policies, guidelines/DMPs...

Plans, budgets, approaches aligned across units/entities
Stakeholder buy-in

OA-OD capacity, support, visibility

■ phase I OA/OD, build in phase II
■ Big Data Platform



Aspiration

Global discovery
Federated search across centers



Categorized
content type



Faceted results

Contents referenced
via standard geo-
coordination (ISO)



Collection of tools

Toolkit for
analytics



Search CERES

drought tolerance maize

Search

Results found in 17 databases for "drought tolerance maize"

Literature

Books	28	books and reports
Policy briefs	0	policy briefs
Training, extension	0	books, journals, and more
AgPub	170	abstracts and citations
AgPub Central	4,580	full-text journal articles

Geographies

Asia	30	resources for sub-regions and countries in Asia
Africa	82	resources for sub-regions and countries in Africa
Middle East	11	resources for countries in the Middle East
Europe	13	resources for sub-regions and countries in Europe
N. America	17	resources for sub-regions and countries in N. America
S. America	25	resources for sub-regions and countries in S. America

Toolkit

Technology tracker	0	technology adoption tracking
Agri-semantics	5	ontologies, vocabularies
AMKN	0	climate change adaptation and mitigation knowledge network
Activity mapper	2	project and activity mapping
Methods	42	methodologies

Subjects

Agroforestry	3	resources related to agroforestry
Agronomy	22	resources related to agronomy
Aquaculture/fisheries	0	resources related to aquaculture/fisheries
Climate change	155	resources related to climate change
GIS/remote sensing	33	resources related to GIS/remote sensing
Genebank	53	genebank resources
Genetic/genomic	5	genetic/genomic resources
Hydrology/water mgmt.	0	resources related to hydrology/water management
Livestock/animal breeding	120	resources related to livestock/animal breeding
Natural resource mgmt.	60	resources related to natural resource management
Plant breeding	134	resources related to plant breeding
Plant protection	13	resources related to plant protection
Socioeconomics/livelihoods	47	resources related to socioeconomics/livelihoods
Other	56	resources related to other subjects

Machine-readable
Human + machine
readable content



Improved access, reuse



High precision, integration
via controlled
vocabularies, ontologies

Search NCBI databases

[Help](#)

drought toler* maize

Search

Results found in 13 databases for "drought toler* maize"

Literature

Books	24	books and reports
MeSH	0	ontology used for PubMed indexing
NLM Catalog	0	books, journals and more in the NLM Collections
PubMed	197	scientific & medical abstracts/citations
PubMed Central	4,058	full-text journal articles

Health

ClinVar	0	human variations of clinical significance
dbGaP	0	genotype/phenotype interaction studies
GTR	0	genetic testing registry
MedGen	0	medical genetics literature and links
OMIM	0	online mendelian inheritance in man
PubMed Health	1	clinical effectiveness, disease and drug reports

Genomes

Assembly	0	genome assembly information
BioProject	10	biological projects providing data to NCBI
BioSample	10	descriptions of biological source materials
Clone	0	genomic and cDNA clones
dbVar	0	genome structural variation studies
Genome	0	genome sequencing projects by organism
GSS	0	genome survey sequences
Nucleotide	952	DNA and RNA sequences
Probe	0	sequence-based probes and primers

Genes

EST	3	expressed sequence tag sequences
Gene	7	collected information about gene loci
GEO DataSets	7	functional genomics studies
GEO Profiles	0	gene expression and molecular abundance profiles
HomoloGene	0	homologous gene sets for selected organisms
PopSet	0	sequence sets from phylogenetic and population studies
UniGene	0	clusters of expressed transcripts

Proteins

Conserved Domains	0	conserved protein domains
Protein	47	protein sequences
Protein Clusters	0	sequence similarity-based protein clusters
Structure	0	experimentally-determined biomolecular structures

Chemicals

BioSystems	26	molecular pathways with links to genes, proteins and chemicals
PubChem BioAssay	0	bioactivity screening studies
PubChem Compound	0	chemical information with structures, information and links
PubChem Substance	0	deposited substance and chemical information

Nucleotide

Nucleotide ▾

drought toler* maize



Search

[Create alert](#) [Advanced](#)[Help](#)

Species

Plants (56)
Bacteria (6)
[Customize ...](#)

Molecule types

genomic DNA/RNA (934)
mRNA (9)
[Customize ...](#)

Source databases

INSDC (GenBank) (945)
RefSeq (7)
[Customize ...](#)

Genetic

compartments

Chloroplast (1)
Plastid (1)

Sequence length

[Custom range...](#)

Release date

[Custom range...](#)

Revision date

[Custom range...](#)

[Clear all](#)[Show additional filters](#)

Summary ▾ 20 per page ▾ Sort by Default order ▾

Send: ▾

Filters: [Manage Filters](#)

Items: 1 to 20 of 952

<< First < Prev Page 1 of 48 Next > Last >>

i Found 955 nucleotide sequences. Nucleotide (952) EST (3)

- ☐ [Zea mays subsp. mays](#) cultivar CIMBL55 chromosome 9 clone BAC VPP1, *** SEQUENCING IN PROGRESS ***, 6 ordered pieces

106,534 bp linear DNA

Accession: KT923087.1 GI: 942528088

[GenBank](#) [FASTA](#) [Graphics](#)

- ☐ [Zea mays](#) putative NAC domain transcription factor superfamily protein (LOC100502408), mRNA

1,705 bp linear mRNA

Accession: NM_001196886.1 GI: 308080781

[GenBank](#) [FASTA](#) [Graphics](#)

- ☐ [Zea mays](#) uncharacterized LOC542228 (cl1856_2(520)), mRNA

3. 4,382 bp linear mRNA

Accession: NM_001111838.1 GI: 162458741

[GenBank](#) [FASTA](#) [Graphics](#)

- ☐ [Zea mays](#) CAAT-box DNA binding protein subunit B (NF-YB) (LOC542390), mRNA

4. 1,102 bp linear mRNA

Accession: NM_001112582.2 GI: 1007362279

[GenBank](#) [FASTA](#) [Graphics](#)

- ☐ [Zea mays](#) serine/threonine protein kinase 2.2 (SnRK2.2) precursor RNA, complete cds

5. 1,604 bp linear pre-RNA

Accession: KU167546.1 GI: 1024277214

[GenBank](#) [FASTA](#) [Graphics](#)

- ☐ [Zea mays](#) cultivar inbred line HKI-335 angustifolia 3 (AN3) mRNA, complete cds

6. 1,227 bp linear mRNA

Results by taxon

Top Organisms [\[Tree\]](#)

[Zea mays](#) (48)
[synthetic construct](#) (14)
[Arabidopsis thaliana](#) (8)
[Pseudomonas lini](#) (1)
[Pseudomonas aeruginosa](#) (1)
[All other taxa](#) (880)

[More...](#)

Find related data

Database: [Select](#)[Find items](#)

Search details

(drought tolerance[All Fields] OR drought tolerant[All Fields]) AND ("Zea mays"[Organism] OR "Zea mays subsp. mays"[Organism] OR maize[All Fields])

[Search](#)[See more...](#)

Recent activity

[Turn Off](#) [Clear](#)

drought toler* AND maize (952)

Nucleotide


drought toler* AND maize AND (refseq[filter]) (7)

Nucleotide

Zea mays uncharacterized LOC542228 (cl1856_2(520)), mRNA

NCBI Reference Sequence: NM_001111838.1


[FASTA](#) [Graphics](#)

[Go to:](#) 

LOCUS NM_001111838 4382 bp mRNA linear PLN 23-APR-2017
DEFINITION Zea mays uncharacterized LOC542228 (cl1856_2(520)), mRNA.
ACCESSION NM_001111838
VERSION NM_001111838.1
KEYWORDS RefSeq.
SOURCE Zea mays
ORGANISM [Zea mays](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD
clade; Panicoideae; Andropogonodae; Andropogoneae; Tripsacinae;
Zea.
REFERENCE 1 (bases 1 to 4382)
AUTHORS Lu Y, Li Y, Zhang J, Xiao Y, Yue Y, Duan L, Zhang M and Li Z.
TITLE Overexpression of Arabidopsis molybdenum cofactor sulfurase gene
confers drought tolerance in maize (Zea mays L.)
JOURNAL PLoS ONE 8 (1), E52126 (2013)
PUBMED [23326325](#)
REMARK GeneRIF: Data indicate that overexpression of Arabidopsis
molybdenum cofactor sulfurase gene (LOS5) in maize markedly
enhanced the expression of ZmA0 and aldehyde oxidase (AO) activity,
leading to abscisic acid (ABA) accumulation and increased drought
tolerance.
REFERENCE 2 (bases 1 to 4382)
AUTHORS Sekimoto H, Seo M, Dohmae N, Takio K, Kamiya Y and Koshihara T.
TITLE Cloning and molecular characterization of plant aldehyde oxidase
J. Biol. Chem. 272 (24), 15280-15285 (1997)
PUBMED [9182554](#)
REFERENCE 3 (bases 1 to 4382)
AUTHORS Koshihara T, Saito E, Ono N, Yamamoto N and Sato M.
TITLE Purification and Properties of Flavin- and Molybdenum-Containing
Aldehyde Oxidase from Coleoptiles of Maize
JOURNAL Plant Physiol. 110 (3), 781-789 (1996)
PUBMED [12226218](#)
COMMENT PROVISIONAL [REFSEQ](#): This record has not yet been subject to final
NCBI review. The reference sequence was derived from [D88451.1](#).

##Evidence-Data-START##

Transcript exon combination :: D88451.1, SRR3147030.20297.1

Customize view 

Analyze this sequence

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

Articles about the cl1856_2(520) gene

Overexpression of Arabidopsis molybdenum
cofactor sulfurase gene confers [PLoS One. 2013]

Sequencing, mapping, and analysis of 27,455
maize full-length cDNAs. [PLoS Genet. 2009]

Anchoring 9,371 maize expressed sequence
tagged unigenes to the bacte [Plant Physiol. 2004]

[See all...](#)

Pathways for the cl1856_2(520) gene

Tryptophan metabolism

Reference sequence information

RefSeq protein product

See the reference protein sequence for indole-3-
acetaldehyde oxidase (NP_001105308.1).

More about the gene cl1856_2(520)

cl1856_2(520) gene

Also Known As:

ZEAMMB73_Zm00001d034387...

KAVSGNLCRCTGYRPIVDLEDLGLNCFWKKGEPAEVSRLPGYNSGAV
CTFPEFLKSEIKSTMKVNDVPIAASGDGWYHPKSIEELHRLFDSSWFDDSSVKIVAS
NTGSGVYKDQDLYDKYIDIKGIPELSVINKNDKAIELGSVVSISKAIEVLSDGNLVFR
KIADHLNKVASPFVRNTATIGGNIMMAQRLPFESDVATVLLAAGSTVTVQVASKRLCF
TLEEFLEQPPCDSRTLLLSIFIPWGS DYVTFETFRAAPRPFNAVS YVNSAFLARTS
GSLIEDICLAFGAYGVDHAIRAKKVEDFLKGKSLSSFVILEAIKLLKDTVSPSEGT
HHEYRVSLAVSFLFSFLSSLANSSSAPSNIDTPNGSYTHETGSNVDS PERHIK VDSND
LPIRSRQEMVFSDEYKPVGKPIKKVGAEIQASGEAVYVDDIPAPKDCLYGAFIYSTHP
HAHVR SINFKSSLASQKVITVITAKDIPSGGENIGSSFLMQGEALFADPIAEFAGQNI
GVVIAETQRYANMAAKQAWVEYSTENLQPPILTIEDAIQRNSYIQIPPFLAPKPVGDY
NKGMAEADHKILSAEVKLESQYYFYMETQAALAI PDEDNCITIYSSTQMPELTQNLIA
RCLGIPFHNVRVISRRVGGGFGGKAMKATHTACACALAAFKLRPVRMYLDRKTDMMIM
AGGRHPMKAKYSVGFKSDGKITALHDLGINAGISPDVSPLMPRAIIGALKKYNWGT
EFDTKVCKTNVSSKSAMRAPGDVQGSFIAEATIEHVASALALDNTVRRKNLHDFESL
EVFYGESAGEASTYSLVSMFDKLALSPEYQHRAAMIEQFNSSNKWKRGISCVPATYE
VNL RPTPGKVSIMNDGSIAVEVGGIEIGQGLWTKVKQMTAFGLGQLCPDGGEC LLDKV
RVIQADT LSLIQGGMTAGSTTSETSCETVRQSCVALVEKLNPIKESLEAKSNTVEWSA
LIAQASMASVNLSAQPYWTPDPSFKSYLNYGAGTSEVEVDILTGATTILRSDLVYDCG
QSLNPAVDLGQIEGCFVQGIGFFTNEDYKTNSDGLVIHDGTWYK IPTVDNIPKEFNV
EMFNSAPDKKRVLSSKASGEPLVLATSVHCAMREAIRAARKEFSVSTSPAKSAVTFQ
MDVPATMPVVKELCGLDVVERYLENSAASAGPNTAKA"

STS

4068..4349
/gene="cl1856_2(520)"
/gene_synonym="A01; GRMZM2G141535; zmAO-1"
/standard_name="PZA02078-70545-B73"
/db_xref="UniSTS:276330"

polyA_site

4382
/gene="cl1856_2(520)"
/gene_synonym="A01; GRMZM2G141535; zmAO-1"
/note="30 A nucleotides"

ORIGIN

```
1  gtgctgtgtt  gtgctgtgct  gcgtgctgtg  gagggggagg  aggagatggg  gaaggaggca
61  ggggcagcgg  agtcgtcgac  ggtgggtgctg  gccgtcaacg  gcaagcgcta  cgaggcggcc
121  ggcgtggctc  cgtccacgctc  gctgctggag  ttcctccgca  cccagacgcc  cgtcagaggc
181  cccaagctcg  gctgcggcga  aggtggctgc  ggtgcatgcg  tggtcctcgt  ctccaagtac
241  gacccggcca  cggacgaggt  gaccgagttc  tctgccagct  cctgcctgac  gctgctccac
301  agcgtggacc  gctgctcagt  gaccaccagc  gagggaatcg  gcaacaccag  ggatggctac
361  caccccgctg  agcagcgcct  ctccggcttc  cacgcctcgc  agtgcggctt  ctgcacaccc
421  ggcattgtgca  tgtccatctt  ctccgccctt  gtcaaggccg  acaacaagtc  cgatcgcccc
```

Zea mays uncharacterized LOC542228 (cl1856_2(520)), mRNA

NCBI Reference Sequence: NM_001111838.1

[FASTA](#) [Graphics](#)

[Go to:](#) ☐

LOCUS NM_001111838 4382 bp mRNA linear PLN 23-APR-2017
DEFINITION Zea mays uncharacterized LOC542228 (cl1856_2(520)), mRNA.
ACCESSION NM_001111838
VERSION NM_001111838.1
KEYWORDS RefSeq.
SOURCE Zea mays
ORGANISM [Zea mays](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD
clade; Panicoideae; Andropogonodae; Andropogoneae; Tripsacinae;
Zea.
REFERENCE 1 (bases 1 to 4382)
AUTHORS Lu Y, Li Y, Zhang J, Xiao Y, Yue Y, Duan L, Zhang M and Li Z.
TITLE Overexpression of Arabidopsis molybdenum cofactor sulfurase gene
confers drought tolerance in maize (Zea mays L.)
JOURNAL PLoS ONE 8 (1), E52126 (2013)
PUBMED [23326325](#)
REMARK GeneRIF: Data indicate that overexpression of Arabidopsis
molybdenum cofactor sulfurase gene (LOS5) in maize markedly
enhanced the expression of ZmA0 and aldehyde oxidase (AO) activity,
leading to abscisic acid (ABA) accumulation and increased drought
tolerance.
REFERENCE 2 (bases 1 to 4382)
AUTHORS Sekimoto H, Seo M, Dohmae N, Takio K, Kamiya Y and Koshiba T.
TITLE Cloning and molecular characterization of plant aldehyde oxidase
J. Biol. Chem. 272 (24), 15280-15285 (1997)
PUBMED [9182554](#)
REFERENCE 3 (bases 1 to 4382)
AUTHORS Koshiba T, Saito E, Ono N, Yamamoto N and Sato M.
TITLE Purification and Properties of Flavin- and Molybdenum-Containing
Aldehyde Oxidase from Coleoptiles of Maize
JOURNAL Plant Physiol. 110 (3), 781-789 (1996)
PUBMED [12226218](#)
COMMENT PROVISIONAL [REFSEQ](#): This record has not yet been subject to final
NCBI review. The reference sequence was derived from [D88451.1](#).

##Evidence-Data-START##

Transcript exon combination :: D88451.1, SRR3147030.20297.1

Customize view

Analyze this sequence

[Run BLAST](#)

[Pick Primers](#)

[Highlight Sequence Features](#)

[Find in this Sequence](#)

Articles about the cl1856_2(520) gene

[Overexpression of Arabidopsis molybdenum cofactor sulfurase gene confers drought tolerance in maize \(Zea mays L.\)](#) [PLoS One. 2013]

[Sequencing, mapping, and analysis of 27,455 maize full-length cDNAs.](#) [PLoS Genet. 2009]

[Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial genome](#) [Plant Physiol. 2004]

[See all...](#)

Pathways for the cl1856_2(520) gene

[Tryptophan metabolism](#)

Reference sequence information

[RefSeq protein product](#)

See the reference protein sequence for indole-3-acetaldehyde oxidase (NP_001105308.1).

More about the gene cl1856_2(520)

[cl1856_2\(520\) gene](#)

Also Known As:

[ZEAMMB73_Zm00001d034387...](#)

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected:0

<div>Alignments</div> <div>Download</div> <div>GenBank</div> <div>Graphics</div> <div>Distance tree of results</div>							
	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	Zea mays uncharacterized LOC542228 (cl1856_2(520)), mRNA	8093	8093	100%	0.0	100%	NM_001111838.1
<input type="checkbox"/>	Zea mays CL1856_2 mRNA sequence	7546	7728	96%	0.0	99%	AY109832.1
<input type="checkbox"/>	Sorghum bicolor hypothetical protein, mRNA	6013	6013	93%	0.0	93%	XM_002463717.1
<input type="checkbox"/>	Zea mays full-length cDNA clone ZM_BFc0067H14 mRNA, complete cds	4992	4992	62%	0.0	99%	BT063475.1
<input type="checkbox"/>	PREDICTED: Setaria italica indole-3-acetaldehyde oxidase (LOC101755113), partial mRNA	4724	4724	92%	0.0	88%	XM_014805770.1
<input type="checkbox"/>	PREDICTED: Setaria italica indole-3-acetaldehyde oxidase (LOC101756196), mRNA	4625	4625	92%	0.0	87%	XM_004981431.2
<input type="checkbox"/>	PREDICTED: Setaria italica indole-3-acetaldehyde oxidase (LOC101754702), mRNA	4246	4246	93%	0.0	86%	XM_004981427.3
<input type="checkbox"/>	Sorghum bicolor hypothetical protein, mRNA	4237	4237	92%	0.0	86%	XM_002463715.1
<input type="checkbox"/>	PREDICTED: Setaria italica indole-3-acetaldehyde oxidase-like (LOC101764191), mRNA	4071	4071	91%	0.0	85%	XM_012849241.1
<input type="checkbox"/>	PREDICTED: Zea mays indole-3-acetaldehyde oxidase-like (LOC103644157), mRNA	3975	3975	92%	0.0	84%	XM_008667350.2
<input type="checkbox"/>	Sorghum bicolor hypothetical protein, mRNA	3960	3960	92%	0.0	84%	XM_002463714.1
<input type="checkbox"/>	PREDICTED: Zea mays indole-3-acetaldehyde oxidase-like (LOC103644153), mRNA	3186	3186	84%	0.0	82%	XM_020546989.1
<input type="checkbox"/>	Oryza sativa Japonica Group cDNA clone:J023082B09, full insert sequence	2497	2497	91%	0.0	78%	AK100342.1
<input type="checkbox"/>	PREDICTED: Oryza sativa Japonica Group probable aldehyde oxidase 2 (LOC4334382), mRNA	2495	2495	91%	0.0	78%	XM_015774953.1
<input type="checkbox"/>	Zea mays uncharacterized LOC542229 (cl1856_1), mRNA	2460	4072	91%	0.0	85%	NM_001111839.1
<input type="checkbox"/>	PREDICTED: Oryza sativa Japonica Group probable aldehyde oxidase 3 (LOC4334381), mRNA	2401	2401	91%	0.0	78%	XM_015774952.1
<input type="checkbox"/>	Oryza sativa Japonica Group cDNA, clone: J075146F08, full insert sequence	2401	2401	91%	0.0	78%	AK242113.1
<input type="checkbox"/>	Zea mays PCO077114 mRNA sequence	2368	2368	49%	0.0	86%	AY106434.1

REFERENCE 3 (bases 1 to 4382)
AUTHORS Koshiba T, Saito E, Ono N, Yamamoto N and Sato M.
TITLE Purification and Properties of Flavin- and Molybdenum-Containing
Aldehyde Oxidase from Coleoptiles of Maize
JOURNAL Plant Physiol. 110 (3), 781-789 (1996)
PUBMED [12226218](#)
COMMENT PROVISIONAL [REFSEQ](#): This record has not yet been subject to final
NCBI review. The reference sequence was derived from [D88451.1](#).

##Evidence-Data-START##

Transcript exon combination :: D88451.1, SRR3147030.20297.1
[ECO:0000332]

RNAseq introns :: mixed/partial sample support
SAMN00780135, SAMN00780136
[ECO:0000350]

##Evidence-Data-END##

FEATURES
source Location/Qualifiers
1..4382
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:[4577](#)"
/chromosome="1"
/map="1"
[gene](#) 1..4382
/gene="cl1856_2(520)"
/gene_synonym="A01; GRMZM2G141535; zmAO-1"
/note="uncharacterized LOC542228"
/db_xref="GeneID:[542228](#)"
[CDS](#) 46..4122
/gene="cl1856_2(520)"
/gene_synonym="A01; GRMZM2G141535; zmAO-1"
/EC_number="[1.2.3.7](#)"
/note="aldehyde oxidase; IAA oxidase; aldehyde oxidase1"
/codon_start=1
/product="indole-3-acetaldehyde oxidase"
/protein_id="[NP_001105308.1](#)"
/db_xref="GeneID:[542228](#)"
/translation="MGKEAGAAESSTVVLAVNGKRYEAGVAPSTSLLEFLRTQTPVR
GPKLGCGEggCGACVVLVSKYDPATDEVTEFSASSCLTLLHSVDRCSVTTSEGIGNTR
DGYHPVQQRLSGFHASQCGFCTPGMCMSIFSALVKADNKSDRPDPAGFSKITTSEAE
KAVSGNLRCRTGYRPIVDTCFSFASDVLEDLGLNCFWKKGEEPAEVSRLPGYNSGAV
CTFPEFLKSEIKSTMKVNDVPIAASGDGWYHPKSIEELHRLFDSSWFDDSSVKIVAS
NTGSGVYKDQDLVDKYIDIKGIPELSVINKNDKAIELGSVVSISKAIEVLSDGNLVR
KIADHLNKVASPFVRNTATIGGNIMMAQRLPFESDVATVLLAAGSTVTQVASKRLCF
TLEEFLEQPPCDsRtLLLSIFIPewGSDYVTFETfRAAPRPFgNAVSyVNSAFLaRTS
GSLLIEDICLafGAYGDVdHAIRAKKVEDfLKGKSLSSfVILEAIKLLKdTVSPSEGTT

See the reference protein sequence for indole-3-
acetaldehyde oxidase ([NP_001105308.1](#)).

More about the gene [cl1856_2\(520\)](#)

[cl1856_2\(520\) gene](#)

Also Known As:

[ZEAMMB73_Zm00001d034387...](#)

Related information

[Annotated Genomic](#)

[BioSystems](#)

[Full text in PMC](#)

[Functional Class](#)

[Gene](#)

[Protein](#)

[PubMed](#)

[PubMed \(RefSeq\)](#)

[PubMed \(Weighted\)](#)

[SNP](#)

[Taxonomy](#)


[UniGene](#)

LinkOut to external resources

[GenScript:ORF Clones in your selected vector](#)
[[GenScript:ORF Clones in your ...](#)]

Recent activity

[Turn Off](#) [Clear](#)

 [Zea mays uncharacterized LOC542228](#)
([cl1856_2\(520\)](#)), mRNA Nucleotide

Links from Nucleotide

Showing Current items.

cl1856_2(520) uncharacterized LOC542228 [Zea mays]

Gene ID: 542228, updated on 4-Apr-2017

Summary

- Gene symbol cl1856_2(520)
- Gene description uncharacterized LOC542228
- Locus tag ZEAMMB73_Zm00001d034387
- Gene type protein coding
- RefSeq status PROVISIONAL
- Organism [Zea mays](#)
- Lineage Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD clade; Panicoideae; Andropogonodae; Andropogoneae; Tripsacinae; Zea
- Also known as AO1; zmAO-1; GRMZM2G141535

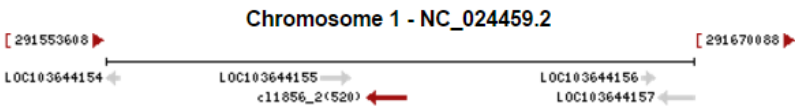
Genomic context

Location: chromosome: 1

See cl1856_2(520) in [Genome Data Viewer](#) [Map Viewer](#)

Exon count: 10

Annotation release	Status	Assembly	Chr	Location
101	current	B73 RefGen_v4 (GCF_000005005.2)	1	NC_024459.2 (291605312..291613050, complement)
100	previous assembly	B73 RefGen_v3 (GCF_000005005.1)	1	NC_024459.1 (286519547..286527285, complement)



Genomic regions, transcripts, and products

Bibliography

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Bibliography
- Variation
- Pathways from BioSystems
- General gene information
 - Markers, Homology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links

Genome Browsers

- Genome Data Viewer
- Map Viewer

Related information

- BioProjects
- BioSystems
- Conserved Domains
- Full text in PMC
- Full text in PMC_nucleotide
- Functional Class
- Gene neighbors
- Genome
- GEO Profiles

Aspiration



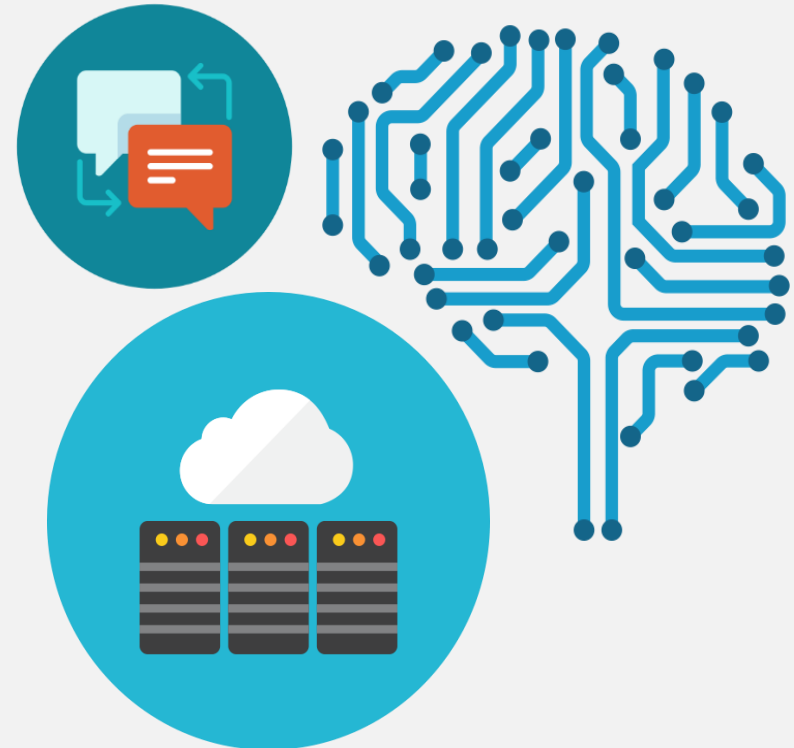
Hey Cigi, when should I plant my maize? How should I manage my crop?

Real-time **decision support** for farmers

Easy **natural language** as an interface

Smart **artificial intelligence** trained by CGIAR and partners

Leveraging multiple open, **harmonized and interoperable** databases





Centro Internacional de Agricultura Tropical
International Center for Tropical Agriculture

Thanks!



INTERNATIONAL
FOOD POLICY
RESEARCH
INSTITUTE

