

ALMAデータアーカイブの概要と アーカイブデータを用いたサイエンス

東京大学大学院理学系研究科
附属天文学教育研究センター
(M2 → D1)
吉村 勇紀



ALMAの概要

▶ 基本的な情報

✓ 観測周波数: 84GHz - 950GHz (およそ波長3mm - 300um)

✓ 結合素子型干渉計

→ 最大50(54)台の12m アンテナ

→ 最大12台の7mアンテナ

✓ 主な特徴

→ 高い感度

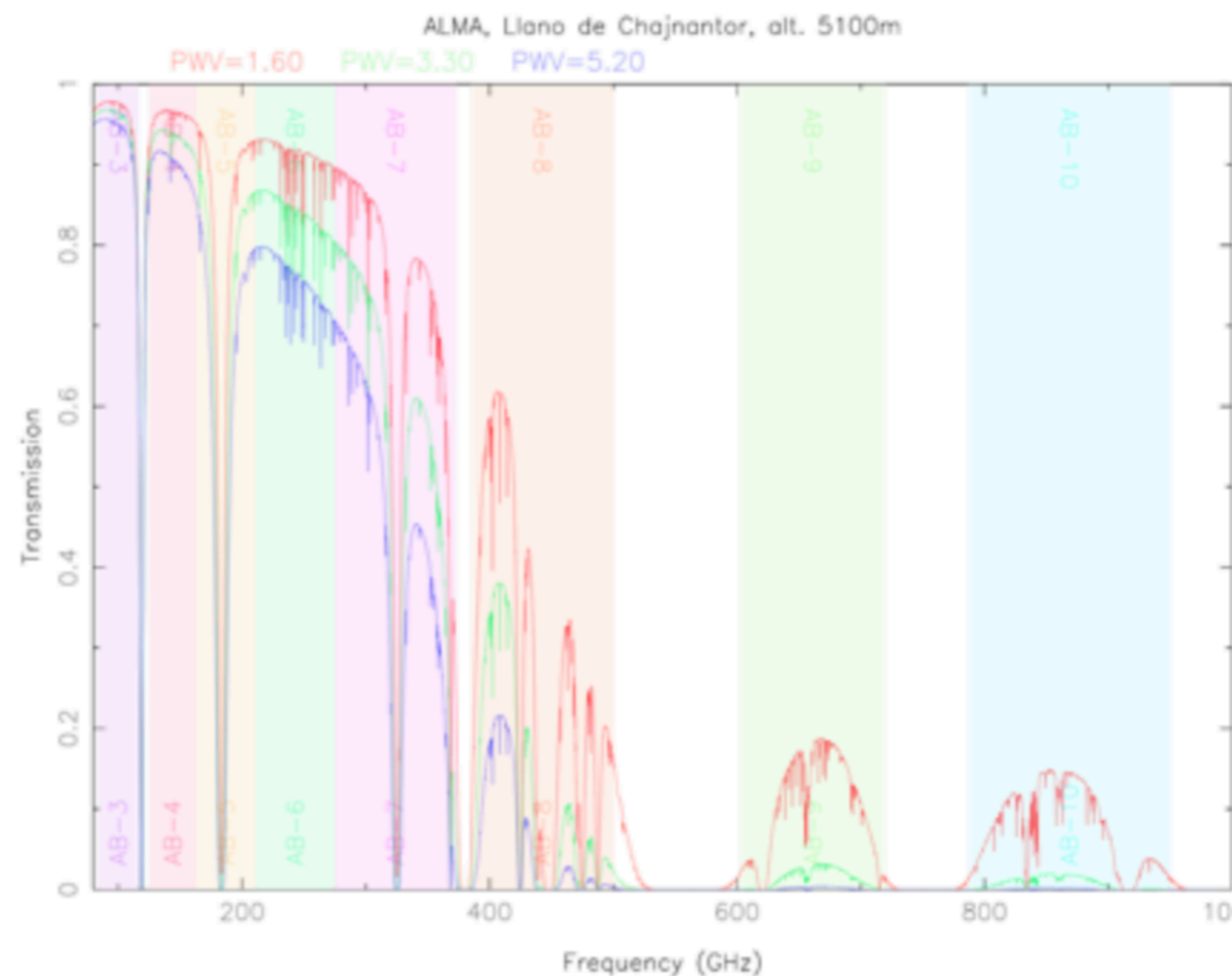
→ 高い角分解能 (現状数十mas程度が達成可能)

→ 狭い視野 (3mmで約1分角; 波長に比例)

▶ ALMAのアーカイブに対するポリシー

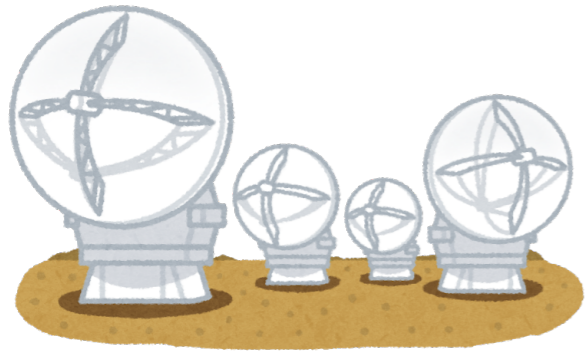
✓ 観測提案PIのアイデアをもとにみんなでより良いアーカイブを作る、という考え方

→ 誰がPIかは全く関係ない



ALMA サイトの大気透過度

ALMA データの概要



観測提案のPI ≠ 観測者

(完全なキュー観測)

生ビジビリティデータ .asdm

一つのプロジェクトあたり、
数十GB ~ 数TB

↓ 各種のキャリブレーション

校正済ビジビリティデータ .ms

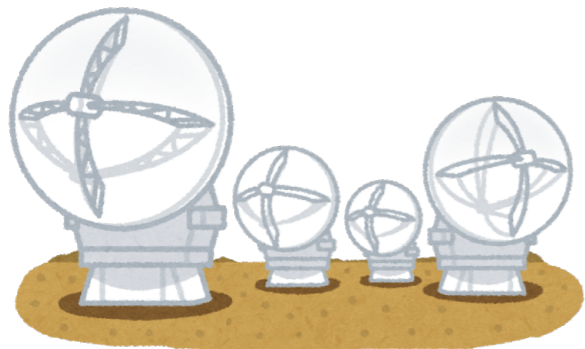
科学観測データ以外にも校正
に必要な様々なメタデータが
付随 (校正天体、気象データ
など)

↓ イメージング
(フーリエ変換、クリーン)

FITS cube .fits

最終生成物: 天球面2次元+周
波数の3次元キューブデータ
→ 可視赤外のIFUに相当

ALMA データの概要



観測提案のPI ≠ 観測者

(完全なキュー観測)

品質保証 (QA)

生ビジビリティデータ .asdm

↓ 各種のキャリブレーション

QA0

観測者がその場で
クイックルックで
最低限の品質を確認

校正済ビジビリティデータ .ms

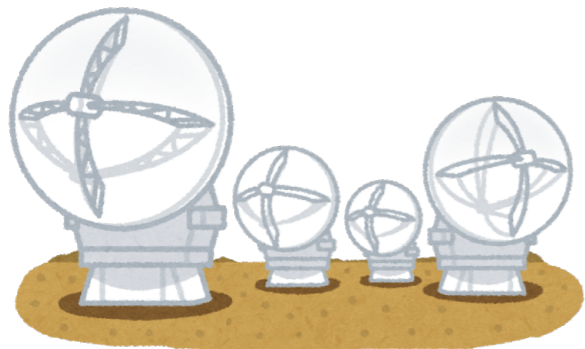
↓ イメージング
(フーリエ変換、クリーン)

QA2

地域センター(ARC)のス
タッフが解析を行い、最
終的なイメージがPIの要
求を満たしているか確認

FITS cube .fits

ALMA データの概要



観測提案のPI ≠ 観測者

(完全なキュー観測)

ALMA science archive

生ビジビリティデータ .asdm

↓ 各種のキャリブレーション →

校正済ビジビリティデータ .ms

↓ イメージング
(フーリエ変換、クリーン) →

FITS cube .fits

→ Science-ready(公称)なのでこれだけでもサイエンスができる！

解析ソフトウェア
(CASA)のスクリプト

scriptForPI.py

必要に応じて部分部分書き換えれば、オリジナルな解析も可能

scriptForImaging.py

ALMA science archiveのインターフェイス

▶ Web query (<https://almascience.nrao.edu/asax/>)

The screenshot displays the ALMA Science Archive interface. On the left, a multi-wavelength image of the galaxy NGC 253 is shown with a yellow circle indicating the field of view (FoV). The coordinates are 00 47 33.300 -25 17 23.00 and the FoV is 18.69'. A white arrow points to the yellow circle with the label "視野の表示" (Field of View display). Below the image, a slider is labeled "多波長画像を表示" (Display multi-wavelength image). On the right, a spectral plot shows the observed frequency range from 100 GHz to 300 GHz. The plot displays several spectral lines with their corresponding molecular species and quantum numbers: CS v=0-2-1, CO v=0-1-0, N2D+ j=2-1, CS v=0-3-2, H13CO+ 2-1, N2H+ v=0-1-2-1, C18O 2-1, CO v=0-2-1, CS v=0-5-4, HCO+ v=0-3-2, N2H+ v=0-1-3-2, H2CO 4(0,4)-3(0,3), CH3OH v=0-6(2,4)-5(1,4), and 13CO v=0-3-2. The plot is labeled "観測周波数範囲、輝線周波数、大気透過度などを表示" (Display observed frequency range, emission line frequency, atmospheric transmission, etc.).

Observations (2) | Projects (3345) | 出版論文 (1797)

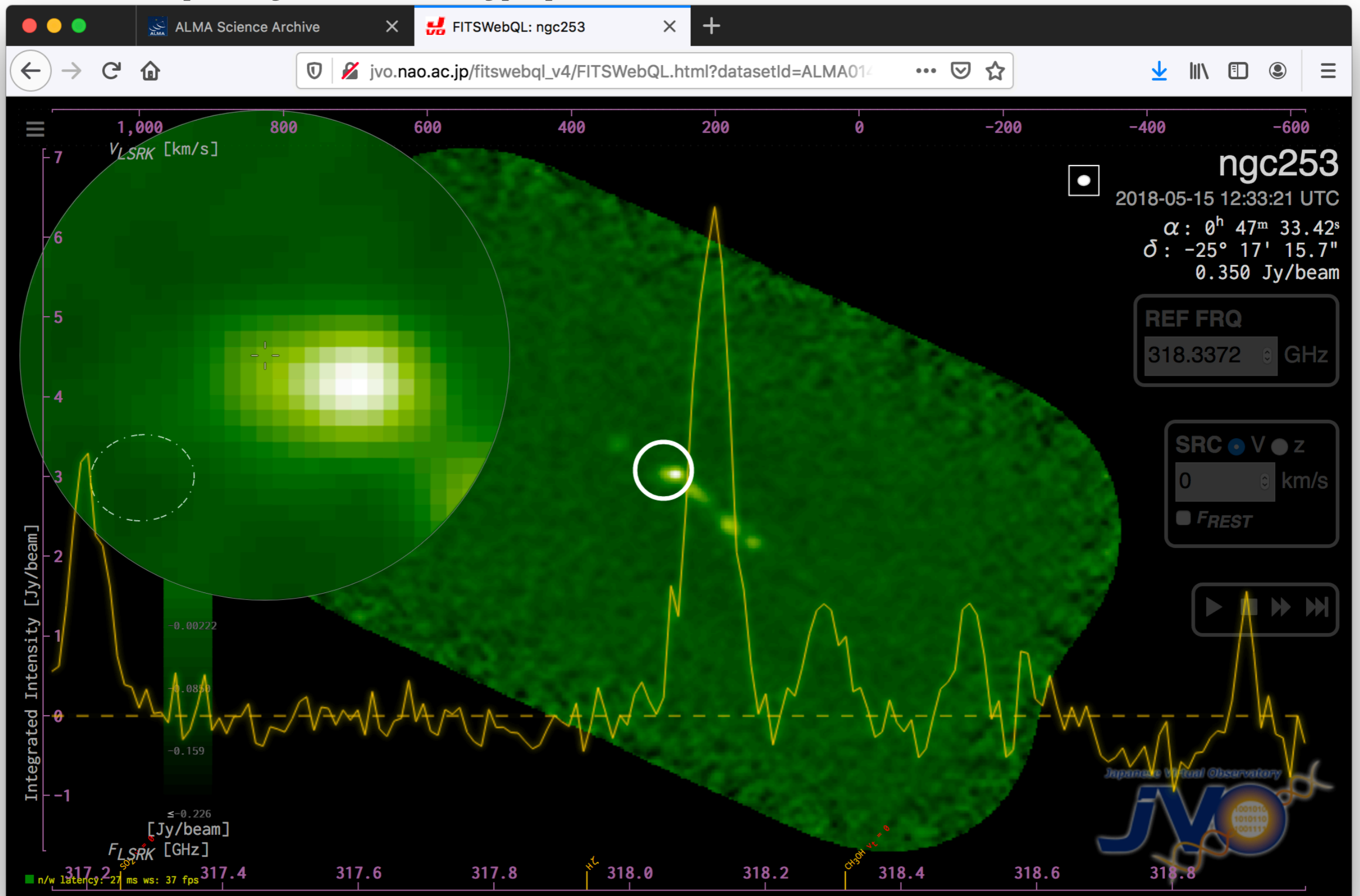
Source: NGC 253 Source is text match only. For source name searches, use the search menu at the top left. Column filters apply only to this tab.

	Project code	Source	Ra	Dec	Band	Cont. sens.	Frequency support	Release date	Publications	Ang. res.
			h:m:s	d:m:s		mJy/beam				arcsec
<input type="checkbox"/>	<input type="text"/>	NGC 253	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
<input type="checkbox"/>	2011.0.00061.S	NGC 253	00:47:33.300	-25:17:23.000	7	0.14	328.36..343.50GHz	2013-05-10	3	1.20
<input type="checkbox"/>	2011.0.00061.S	NGC 253	00:47:33.300	-25:17:23.000	3	0.02	96.37..111.29GHz	2014-01-08	3	1.73

検索結果

ALMA science archiveのインターフェイス

▶ JVO (<https://jvo.nao.ac.jp/portal/alma/archive.do>)



ALMA science archiveのインターフェイス

▶ astroquery (python module)

✓ 様々な天文学上のデータベースにアクセスできる。

✓ ALMAにも対応

→ 検索だけでなく データダウンロードもできる

→ 同じ干渉計でもVLAは検索のみ

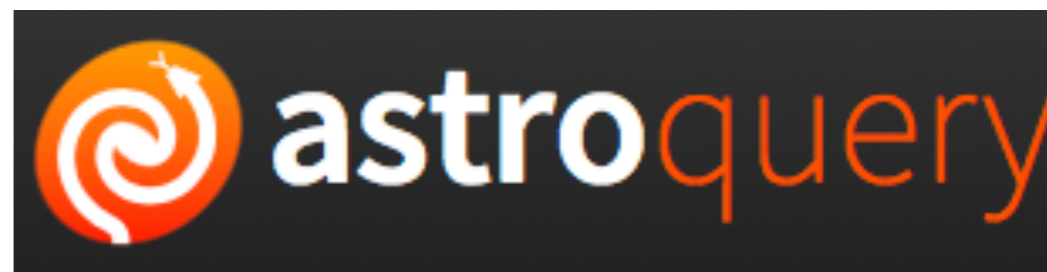
→ 例えば、スクリプトベースの大規模なデータマイニングが可能（後述）

✓ 現状の傾向

→ NRAO主導(?)のためか、電波系、米国系が多く対応している印象。

日本の望遠鏡・サーベイ等も対応を検討してはいかがでしょうか？

<https://astroquery.readthedocs.io/en/latest/>



Available Services

If you're new to Astroquery, a good place to start is the [A Gallery of Queries](#)

- [A Gallery of Queries](#)

The following modules have been completed using a common API:

- [ALMA Queries \(astroquery.alma\)](#)
- [Atomic Line List \(astroquery.atomic\)](#)
- [Besancon Queries \(astroquery.besancon\)](#)
- [Cadc \(astroquery.cadc\)](#)
- [CASDA Queries \(astroquery.casda\)](#)
- [CDS MOC Service \(astroquery.cds\)](#)
- [esa.hubble \(astroquery.esa.hubble\)](#)
- [ESASky Queries \(astroquery.esasky\)](#)
- [ESO Queries \(astroquery.eso\)](#)
- [Gaia TAP+ \(astroquery.gaia\)](#)
- [GAMA Queries \(astroquery.gama\)](#)
- [Gemini Queries \(astroquery.gemini\)](#)
- [HEASARC Queries \(astroquery.heasarc\)](#)
- [HITRAN Queries \(astroquery.hitran\)](#)
- [IRSA Image Server program interface \(IBE\) Queries \(astroquery.ibe\)](#)
- [IRSA Queries \(astroquery.irsa\)](#)
- [IRSA Dust Extinction Service Queries \(astroquery.irsa_dust\)](#)
- [JPL Spectroscopy Queries \(astroquery.jplspec\)](#)
- [MAGPIS Queries \(astroquery.magpis\)](#)
- [MAST Queries \(astroquery.mast\)](#)
- [Minor Planet Center Queries \(astroquery.mpc/astroquery.solarsys\)](#)

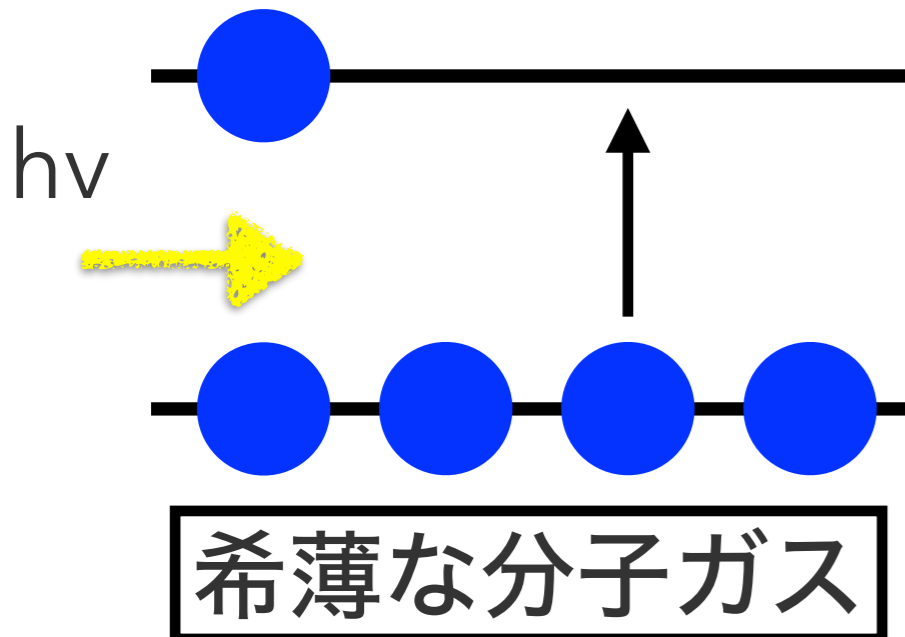
ALMAアーカイブデータを用いた サイエンスの紹介

分子吸収線系の観測に基づく
希薄な分子ガス雲の物理化学的性質の研究

- ① ALMAによる遠方分子吸収線系の高感度観測
- ② ALMA較正用データに対する分子吸収線系探査

希薄な星間分子ガスを観測する手段

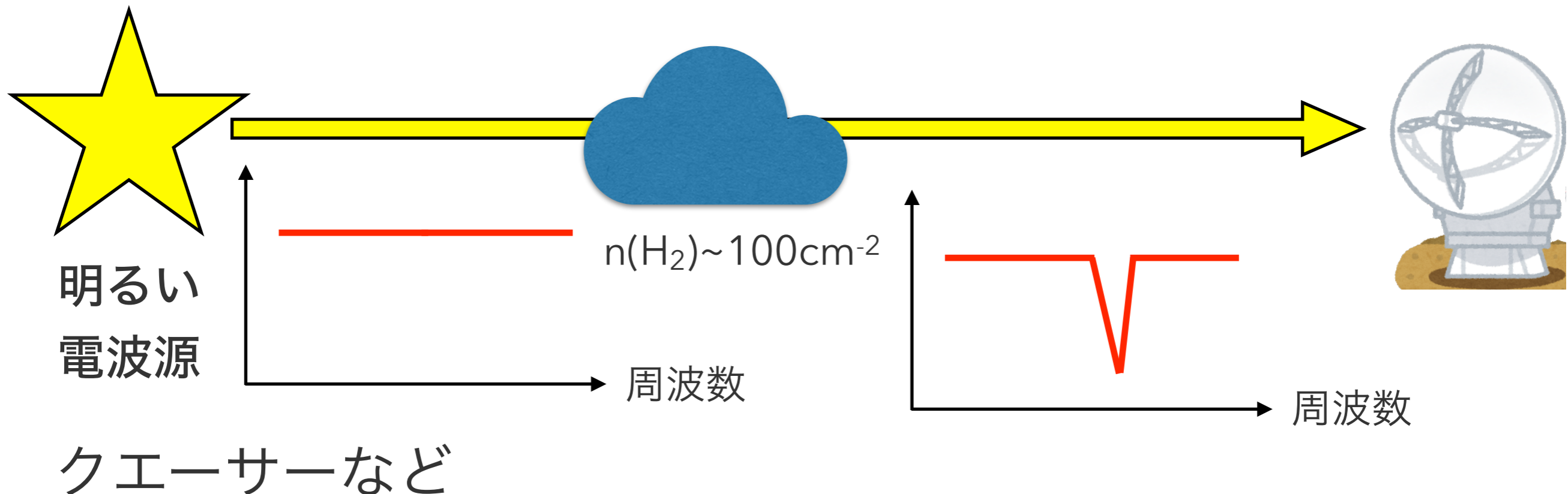
量子的なエネルギー準位



▶ 特徴：背景光源が必要である

✓ **高い感度**で背景光源の連続波スペクトルを
観測する必要がある。

✓ これまで知られている分子吸収線系の数は
極めて少ない。



① ALMAによる遠方分子吸収線系の高感度観測

▶ ALMA観測 (cycle 6)

✓ PI: Y. Yoshimura

▶ 観測天体: B0218+357 (z=0.68)

✓ 約60億光年遠方の分子吸収線系

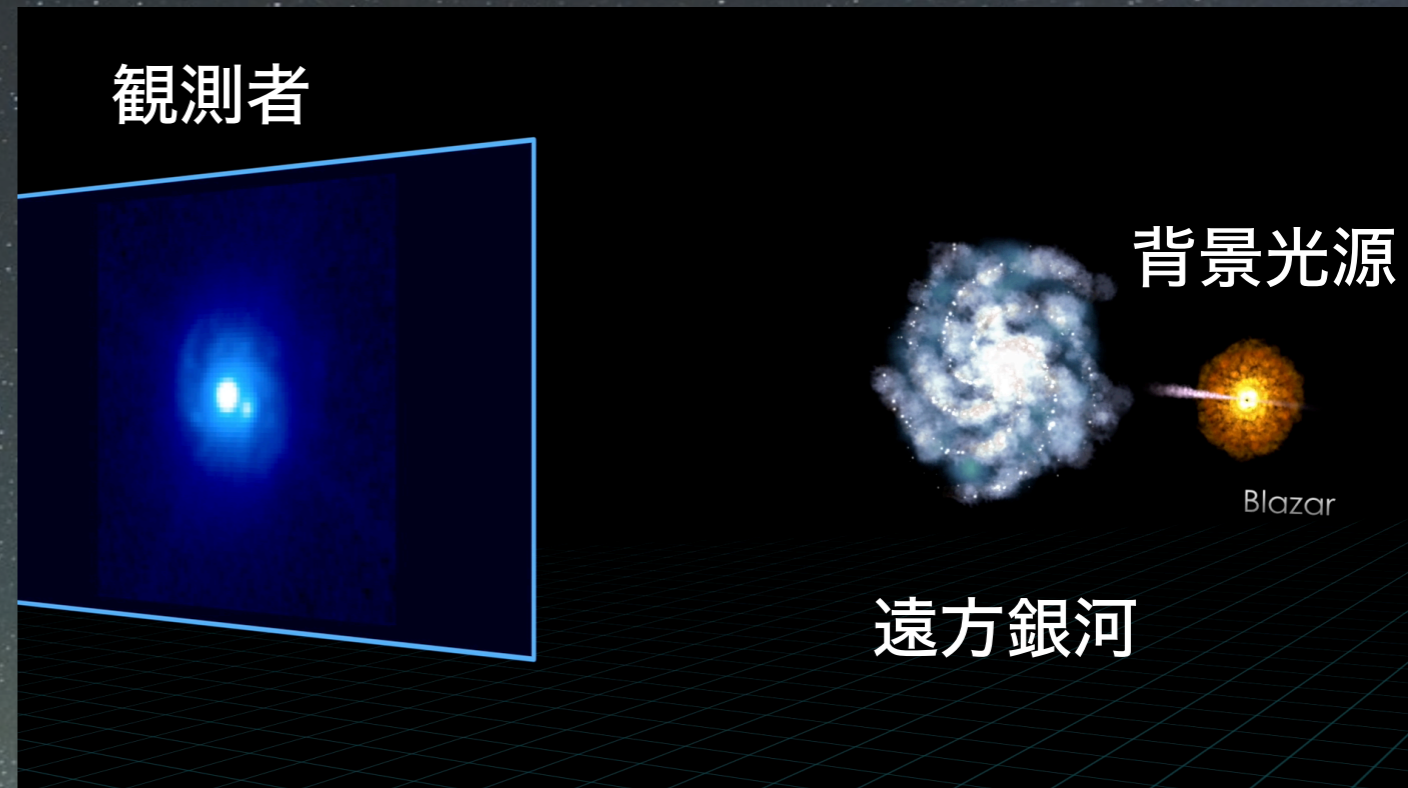
▶ 観測した分子

✓ 水素化リチウム (LiH)

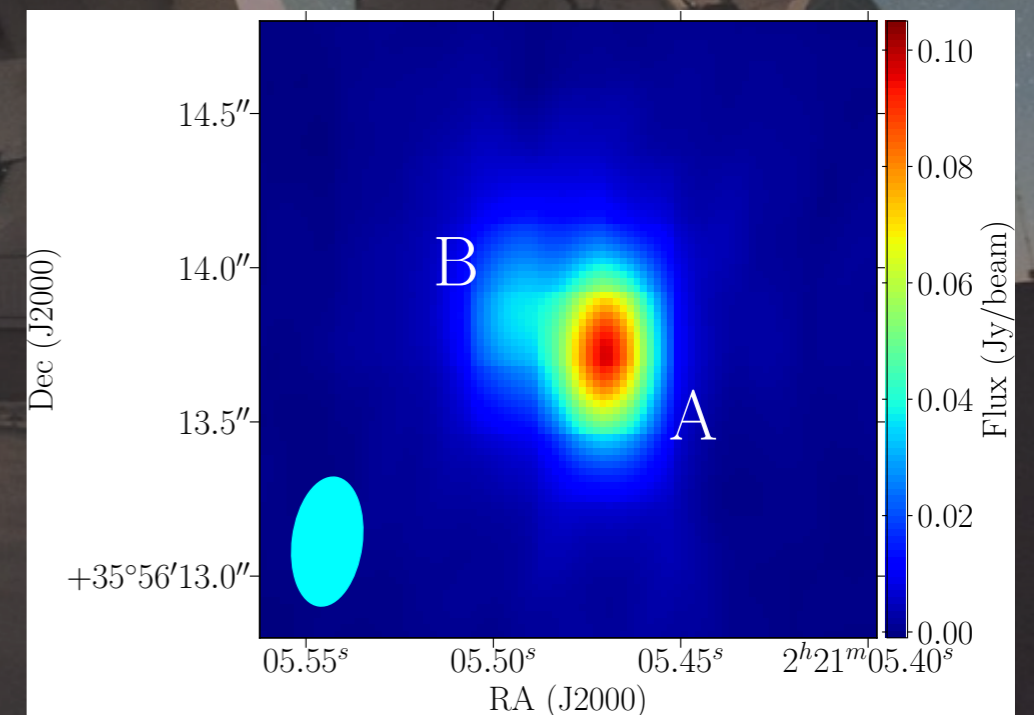
→ 宇宙論におけるリチウム問題解決に向けた新たなプローブとして注目

✓ 硫化水素 (H₂S)

→ 星間空間で豊富な水素化合物のひとつ。



Band 6 continuum image



① ALMAによる遠方分子吸収線系の高感度観測

▶ ALMA観測 (cycle 6)

✓ PI: Y. Yoshimura

▶ 観測天体: B0218+357 (z=0.68)

✓ 約60億光年遠方の分子吸収線系

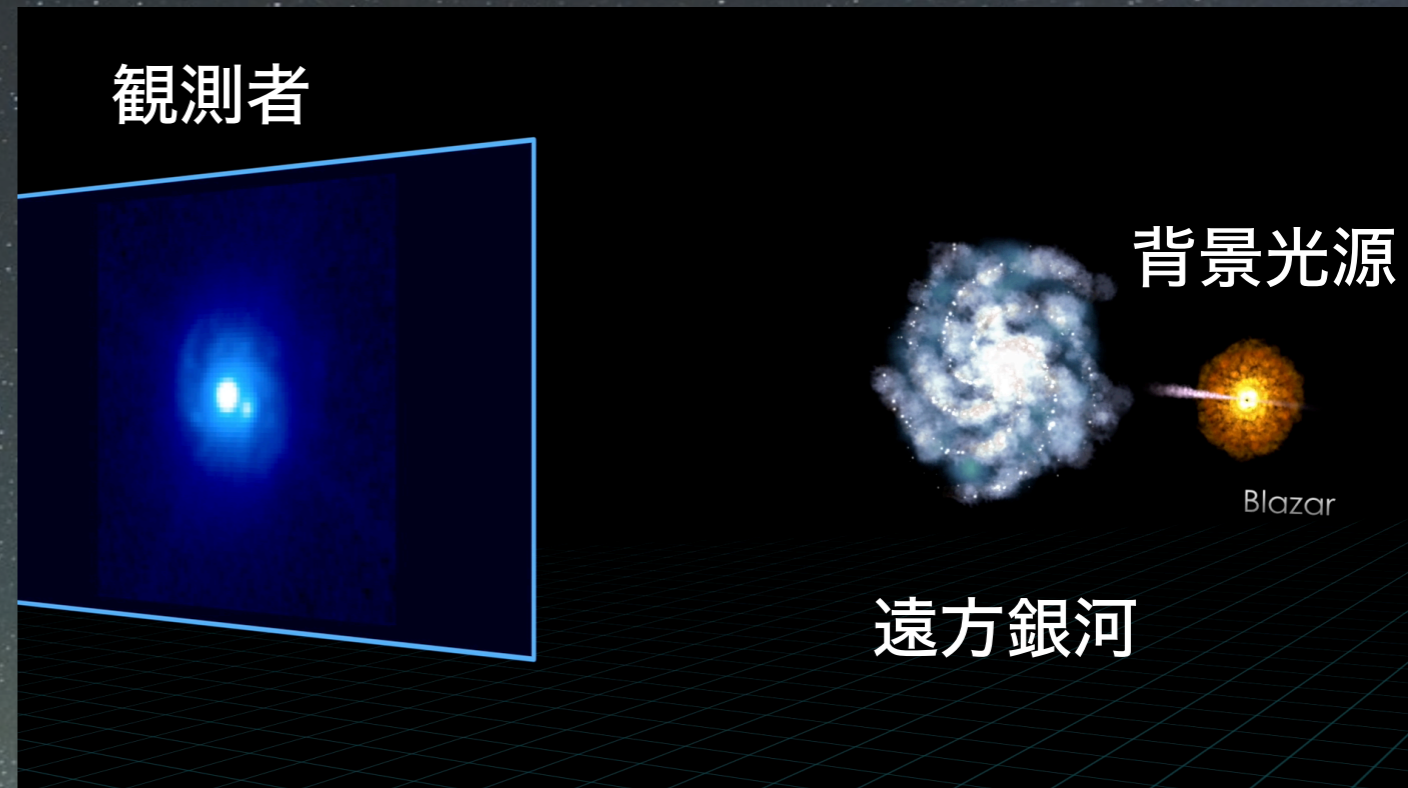
▶ 観測した分子

✓ 水素化リチウム (LiH)

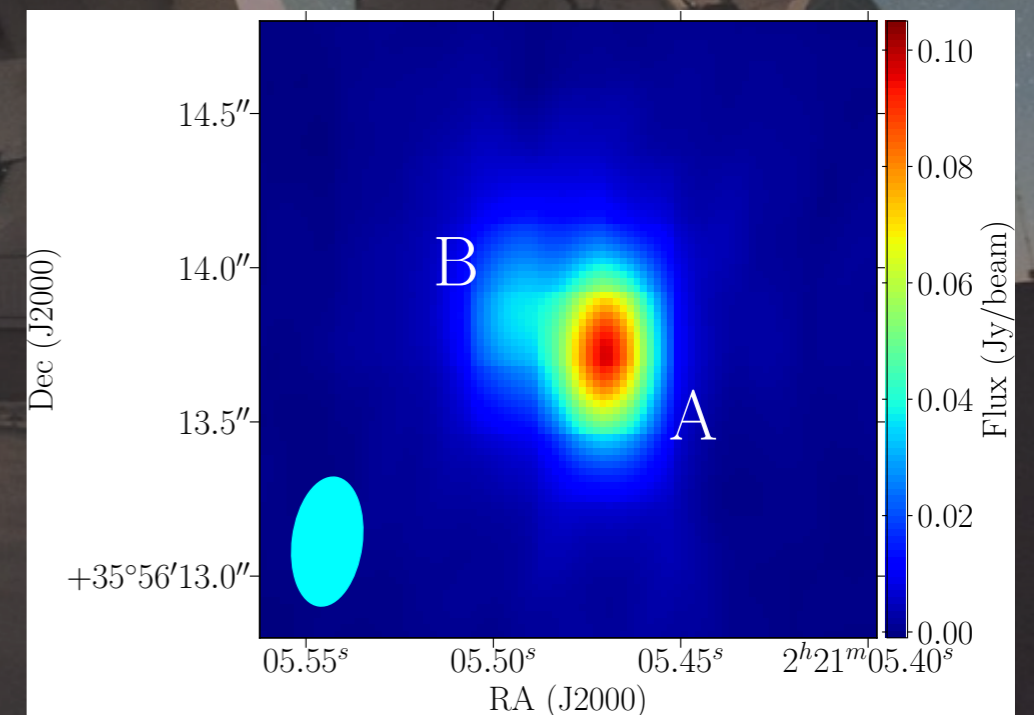
→ 宇宙論におけるリチウム問題解決に向けた新たなプローブとして注目

✓ 硫化水素 (H₂S)

→ 星間空間で豊富な水素化合物のひとつ。



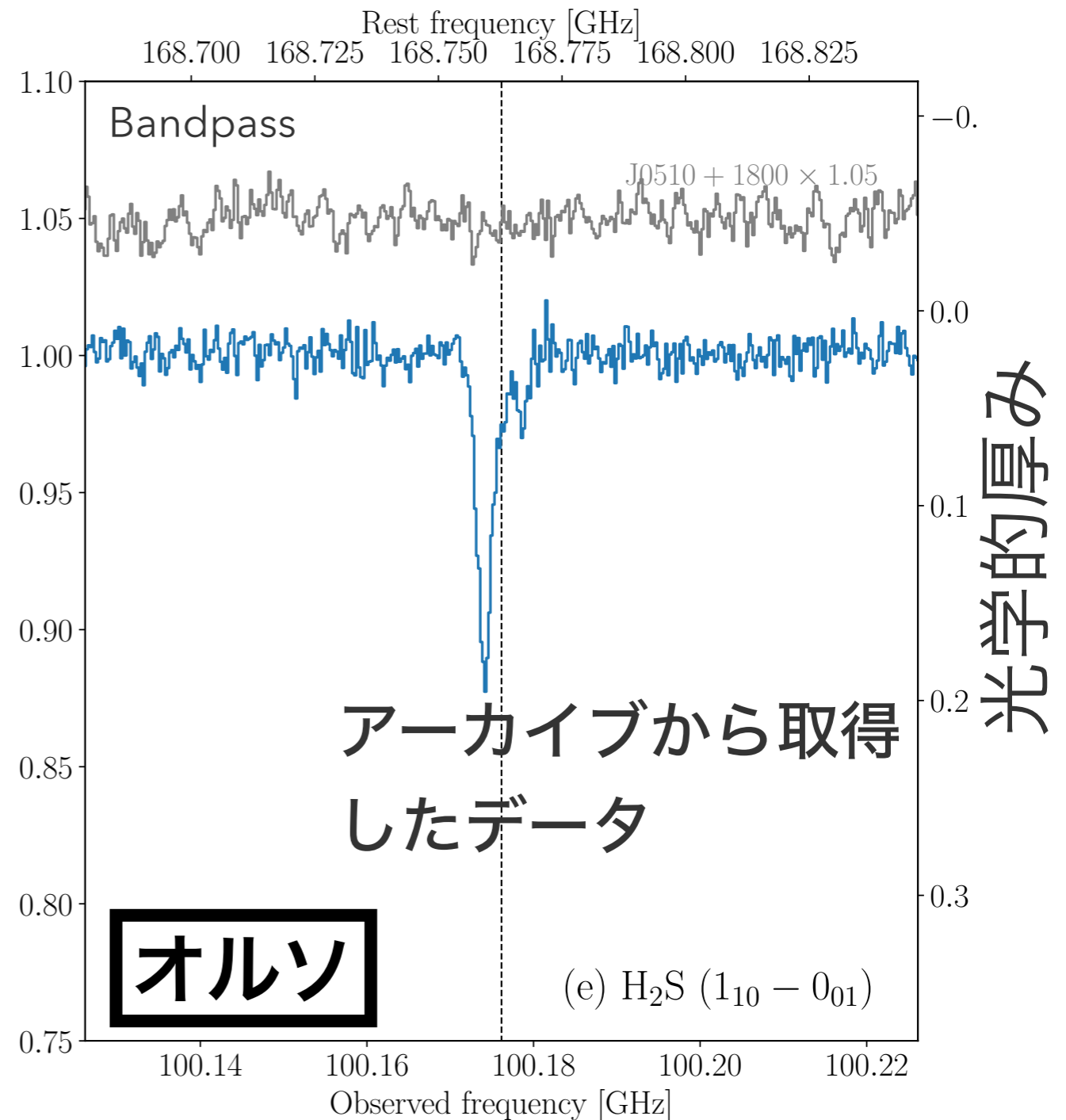
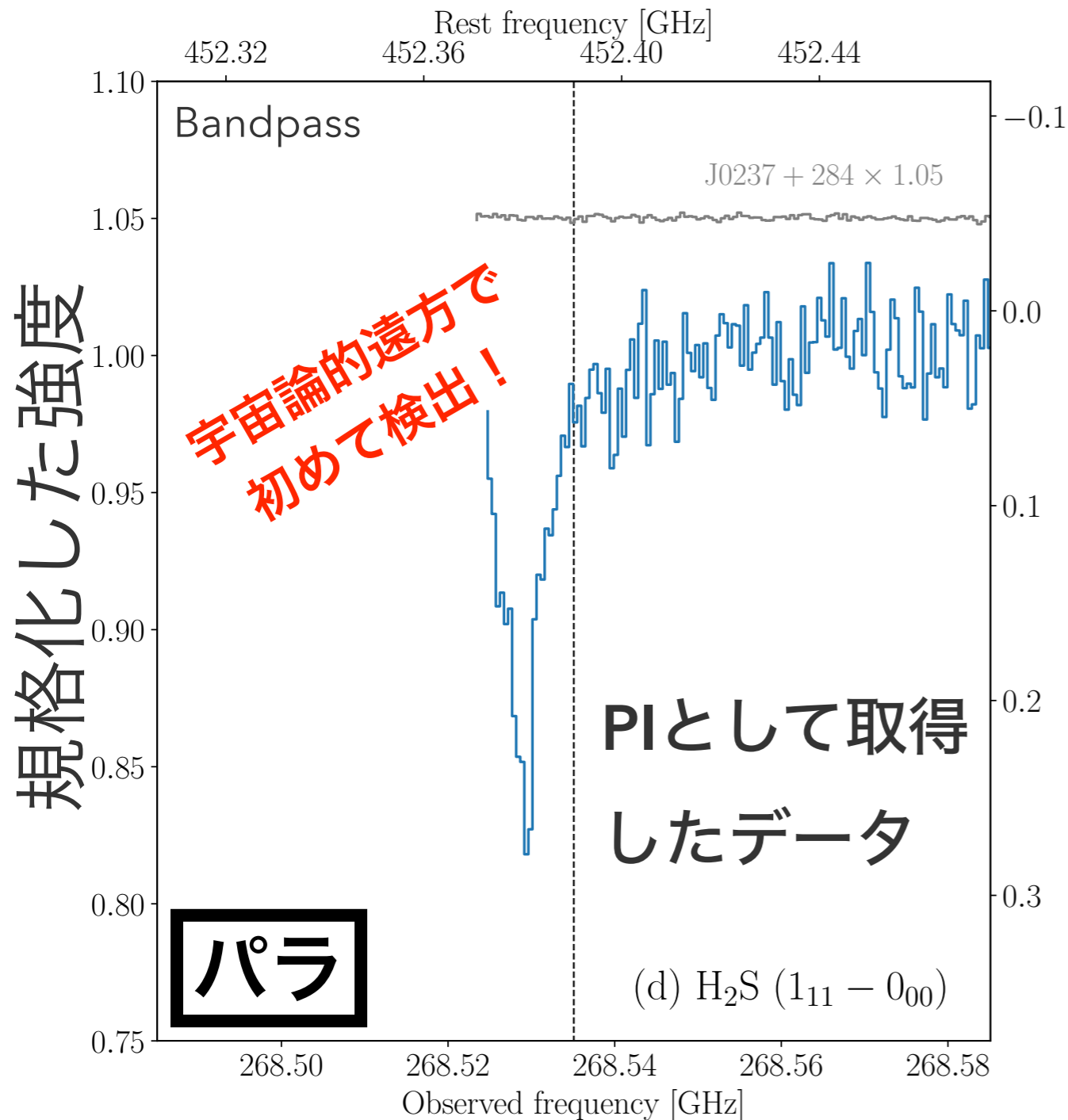
Band 6 continuum image



観測結果: 硫化水素(H₂S)の検出

▶ オルソ・パラ状態それぞれの分子吸収線を検出

✓ 遠方銀河の硫化水素オルソ/パラ比に初めて制限を与える (Yoshimura+, submitted to ApJ)



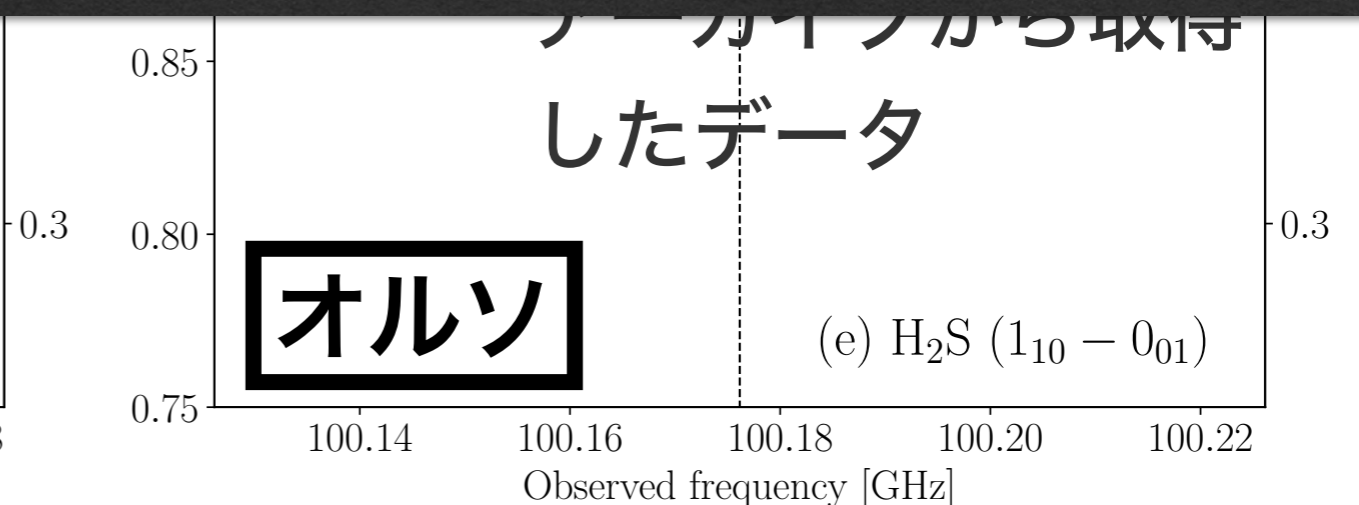
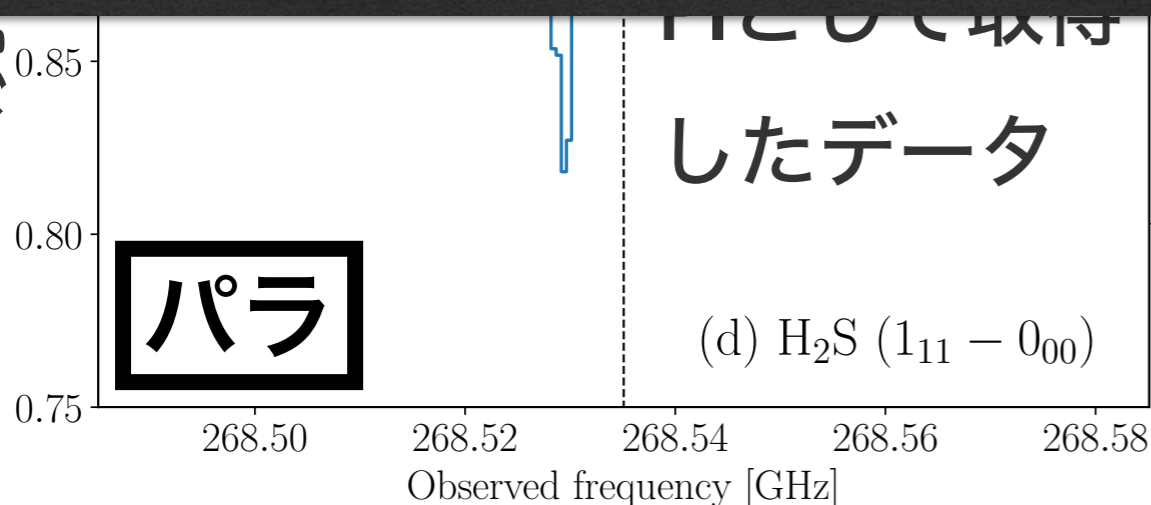
観測結果: 硫化水素(H₂S)の検出

▶ オルソ・パラ状態それぞれの分子吸収線を検出

✓ 遠方銀河の硫化水素オルソ/パラ比に初めて制限を与える (Yoshimura+, submitted to ApJ)



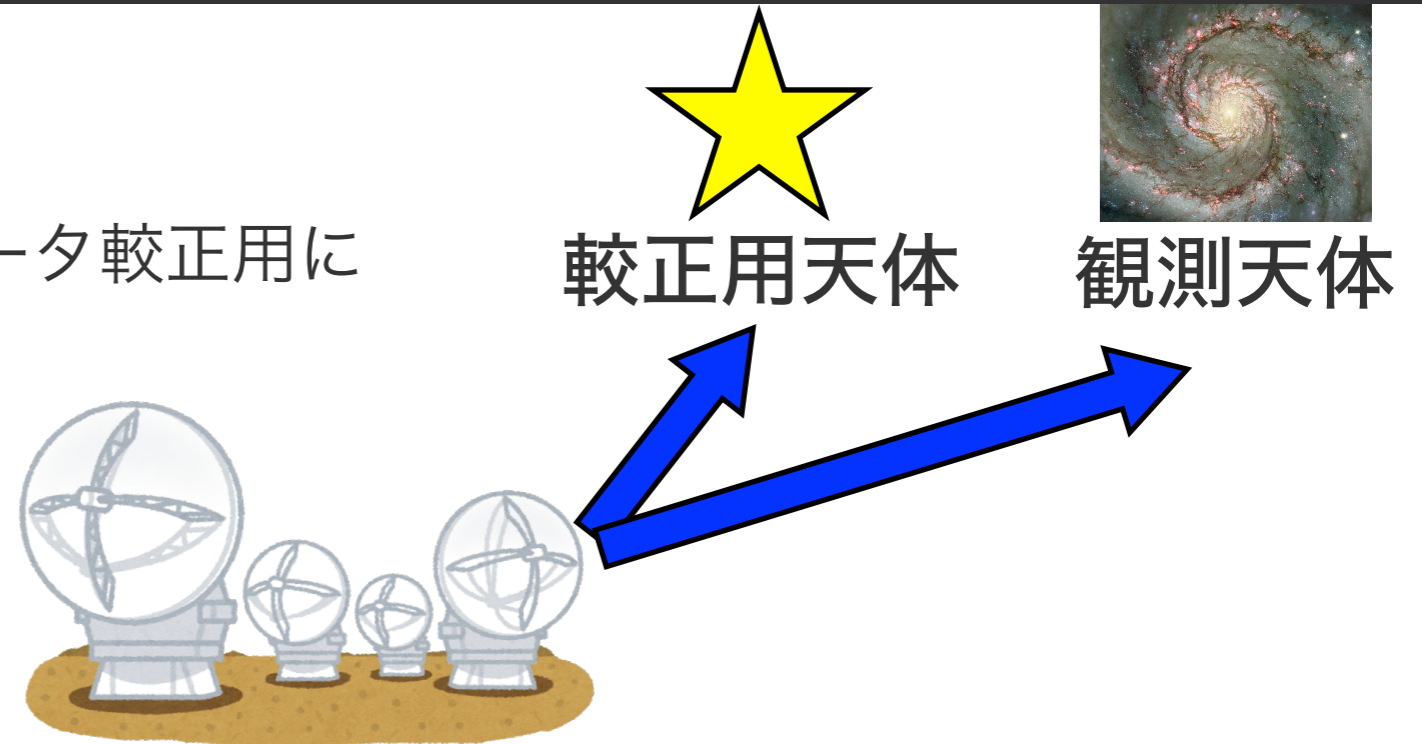
アーカイブデータも組み合わせることで、PIデータのよりよいアウトプットができる



② ALMA校正用データに対する分子吸収線系探査

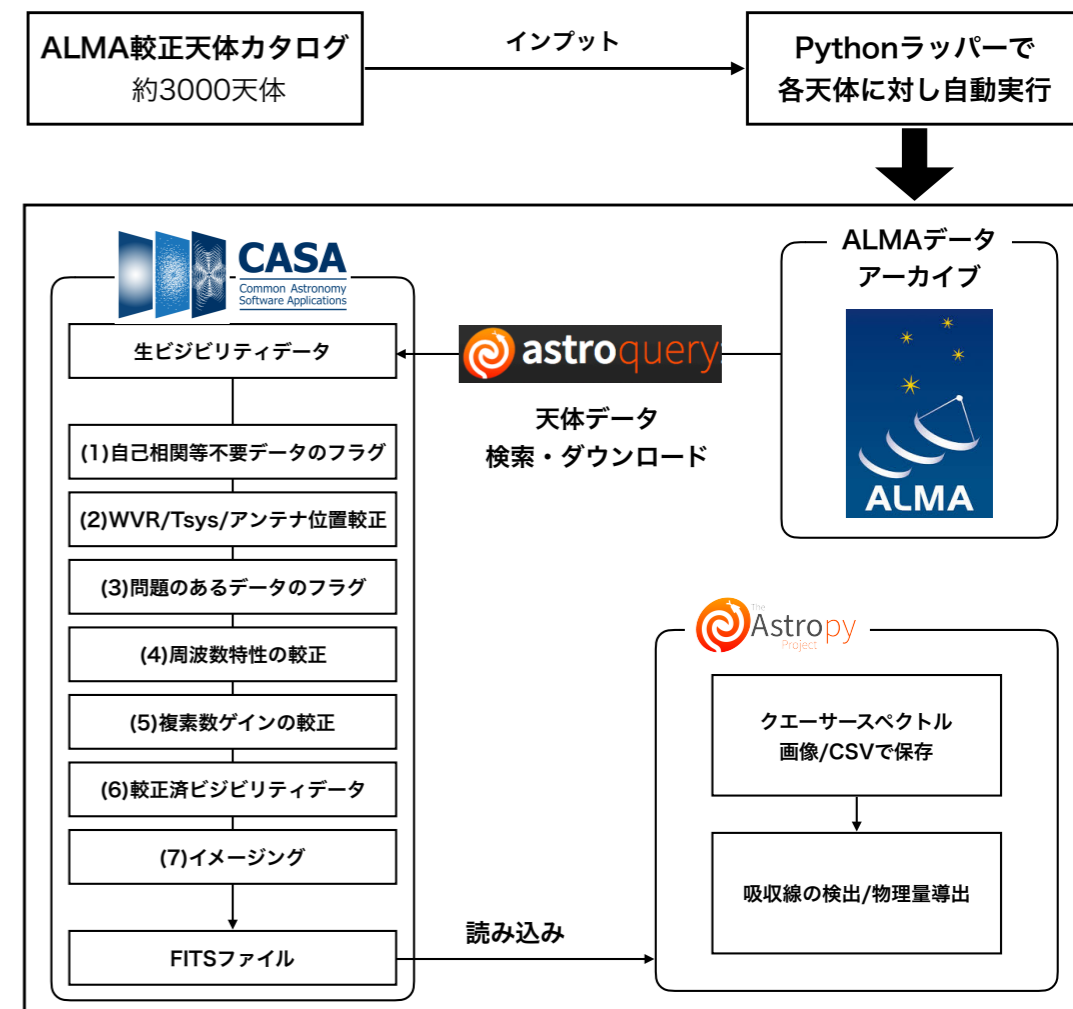
▶ ALMA望遠鏡ビッグデータ

- ✓ ALMAは科学目標天体とは別にデータ校正用に
頻繁にクエーサーを観測
- ✓ 分子吸収線系の潜在的な背景光源

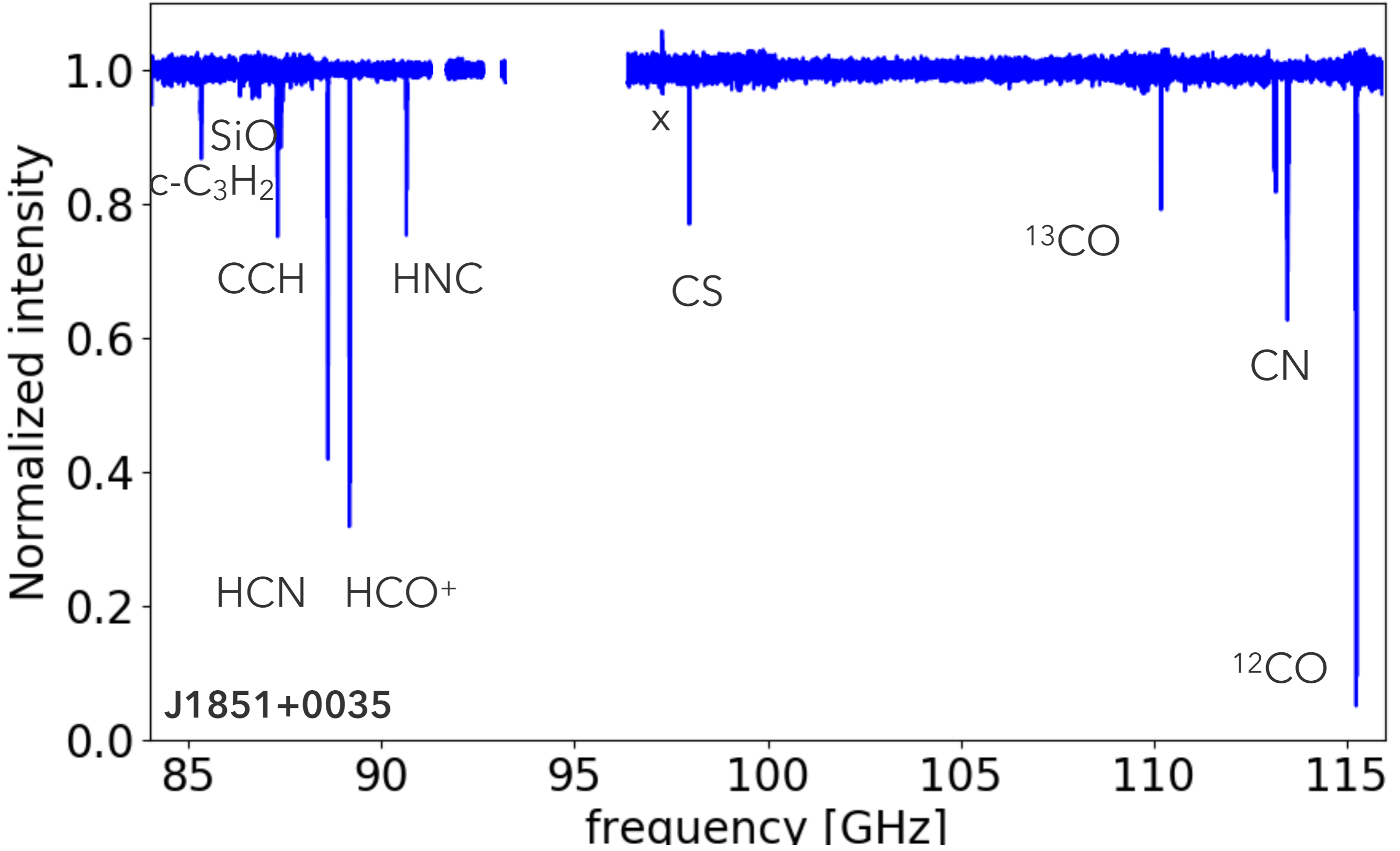


▶ データ取得解析の自動化

- ✓ データの検索・ダウンロード・解析をすべて自動化する**独自のパイプラインを構築**し、悉皆的な分子吸収線系探査を実施
- ✓ **702天体**の2000個超のデータセットを解析

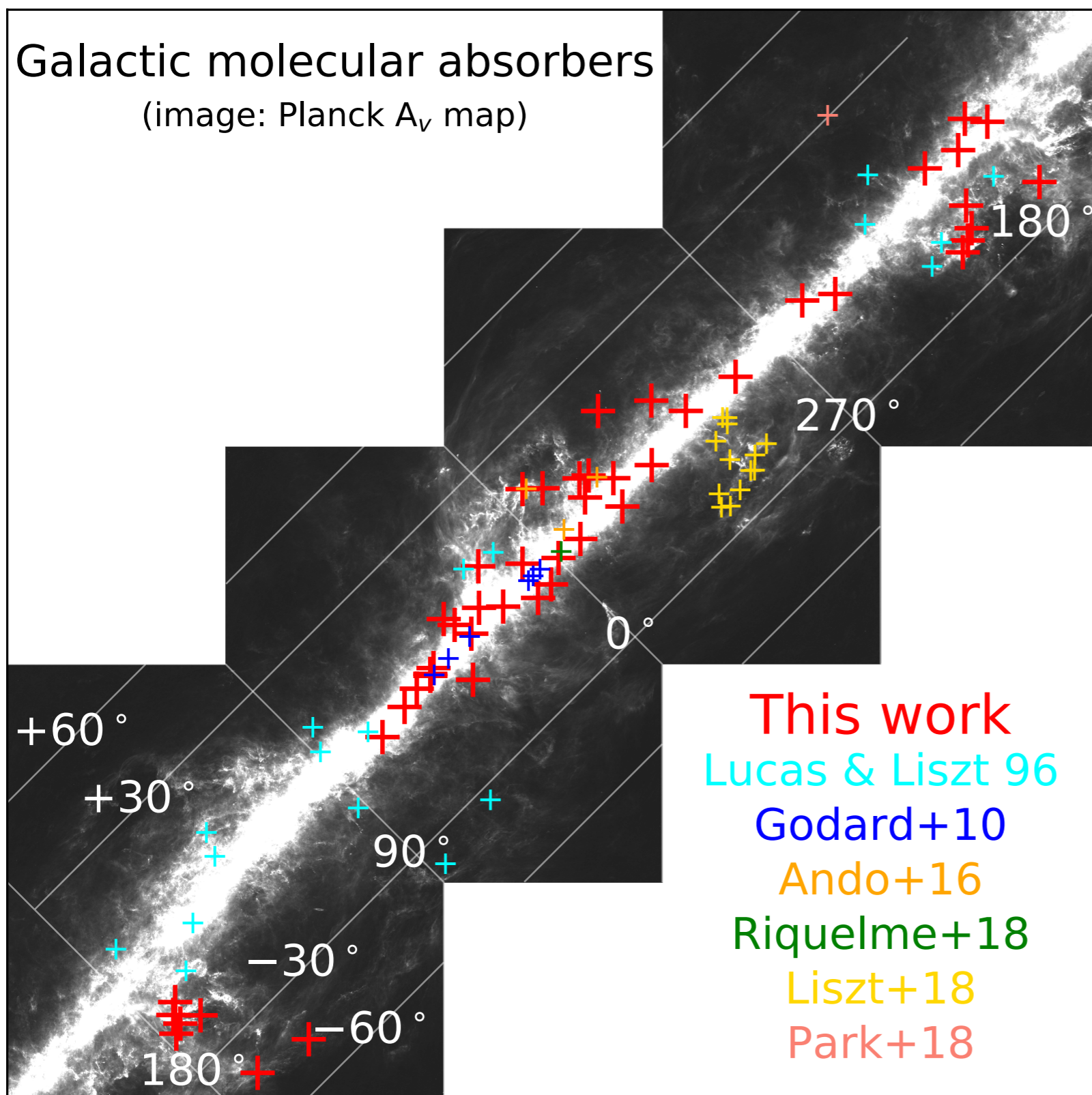


結果：得られた吸収線スペクトルの例



▶ 較正用データのみで、高品質なスペクトルデータが得られた。

多数の銀河系内分子吸収線系を新たに発見



▶ 未発見の銀河系内分子吸収線系を45個発見

▶ 希薄な星間ガスで幾つか興味深い(予想外の)分子の検出

✓ CH₃OH (希薄な分子ガスで**2例目**)

→ 希薄な星間ガスで既に有機分子が形成される可能性。

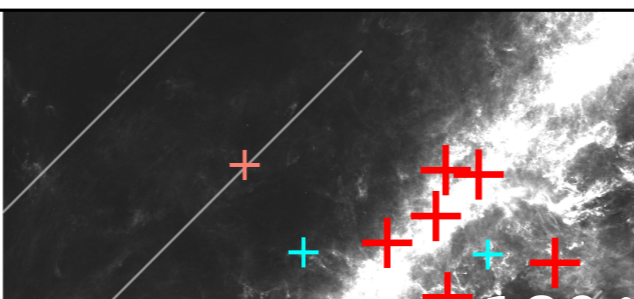
✓ N₂H⁺ (希薄な分子ガスで**初めて**)

→ 局所的な高い宇宙線密度を反映している可能性。

Yoshimura+ in prep.

多数の銀河系内分子吸収線系を新たに発見

Galactic molecular absorbers
(image: Planck A_v map)



▶ 未発見の銀河系内分子吸収線系を45個発見

アーカイブデータは（校正用天体ですら）
宝の山!



や

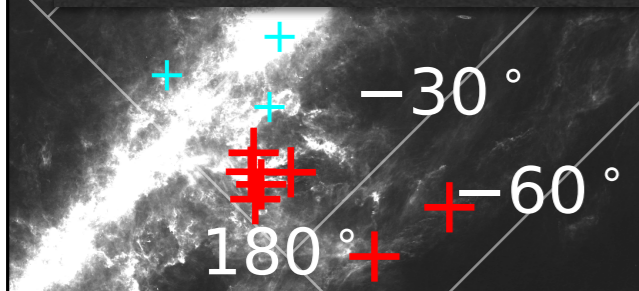


にも期待

Riquelme+18
Liszt+18
Park+18

ている可能性。

Yoshimura+ in prep.



まとめと

アーカイブデータに対する意見など

ALMAのアーカイブシステムについて思うこと

▶ 全般：個人的にはあまり文句がない

▶ ALMAアーカイブシステムのここがすごい

✓ とにかくqueryが使いやすくて、充実している。

→ 新規参入者にはJVO、玄人向けにはastroqueryまで

→ astroqueryは是非もっと多くのアーカイブデータベースにも対応してほしい。

✓ 全データに対して品質保証を行っている。

→ ARCの方々の多大なる労力のおかげ。

✓ メタデータも付随した自己完結したデータ形式(.ms形式)

✓ データの置かれ方（生データ+解析スクリプト+最終生成物）

→ 再現性が高い(というより、ほぼ100%再現できる)

→ 論文化の際も、個々人が解析に用いたスクリプトをgithubなどで公開できると良い？

ALMAのアーカイブシステムについて思うこと

▶ 全般：個人的にはあまり文句がない

▶ **ALMAアーカイブシステムのちょっと不満な点**

✓ データ解析に多大な計算機リソースが必要（メモリ、ストレージ）

→ >100GBのメモリが必要になるケースもある。

→ 恐らく解析ソフトの省メモリ化が必要（CASAのupdateを！）

→ ADCのような共同利用計算機リソースは極めて重要（大変お世話になっています！）

✓ データダウンロード速度

→ データマイニング的な解析では、ここが一番のボトルネックになっている。

→ (例えば) HTTPより高速なプロトコルの導入？

アーカイブデータに求められそうなこと(ALMA等を踏まえて)

▶ 効率的なアーカイブシステムとキュー観測は表裏一体

✓ ALMAはキュー観測だからこそ、各データが自己完結した状態で効率よく活用できる
(決まった観測手順、決まった校正手順)。

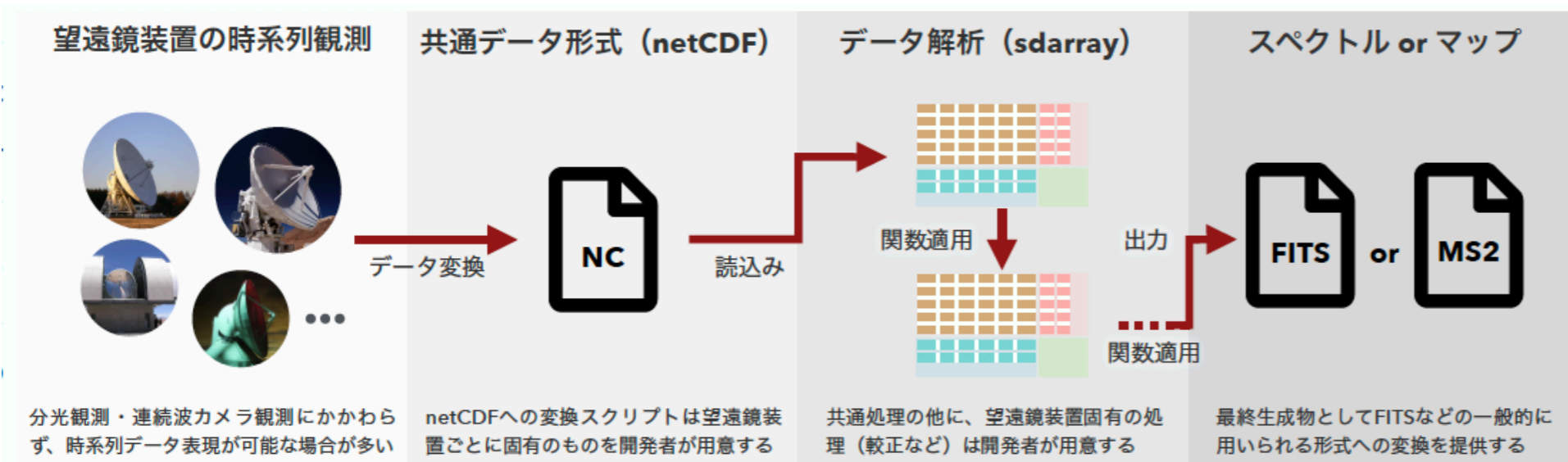
▶ 解析ソフトウェアとデータ形式の統合

✓ ALMA以外の電波望遠鏡もCASA及び.ms形式に対応している(例えば、VLA)

✓ 野辺山45mやASTEのデータも.ms形式に変換し、CASAで解析できる。

✓ 新しい装置/望遠鏡計画にとって、共通データ形式/解析ソフトウェアの議論は必須

→ ビッグデータ天文学の現代においては、省メモリ/ストレージ化を意識する必要がある。



← 大学主導の電波域の新しい装置計画についても、共通データ形式の議論が始まっている (sdarray; 谷口暁星 incl. 吉村他2020天文学会春季年会b講演ポスターより)。

まとめ

▶ アーカイブや公開データに必要なとされる要件

- ✓ 様々なニーズに応じた、使いやすいインターフェイス
- ✓ 較正・解析に必要なメタデータが付随した、自己完結した統一データ形式
 - キュー観測モードが重要
- ✓ 解析スクリプトも付いていると、再現性が高い。
- ✓ 品質保証がされているとなお良い（マンパワーが必要だが）。

▶ アーカイブによりあげられた成果、期待できる成果

- ✓ PIデータだけではなく、アーカイブデータも組み合わせることで、より良い成果
 - 物理量などにさらなる制限がつく。
- ✓ アーカイブの較正用天体を大量に解析することで、サイエンスに発展
 - 次世代望遠鏡についても同様の研究が期待される。