Hubble Legacy Archive (HLA)
Status and Plans

Stefano Casertano
for the HLA Team

HLA Status and News

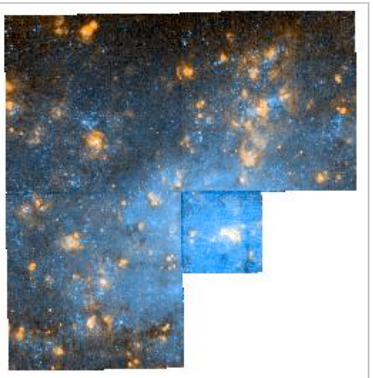
- New organizational presence
  - With the successful Senior Review last summer, HLA is now integrated in MAST
  - STScI renewing focus on Archives as an integral part of the Institute's mission
  - HLA-specific work will be coordinated with the HST Mission Office and the Instrument Teams
- Two major and one intermediate data releases:
  - DR 2 (September 2008):
  - DR 2.5 (December 2008): Contributed High Level Science Products fully integrated in HLA
  - DR 3 (May 2009): NICMOS images, ACS grisms, WFPC2 source lists, prototype mosaics
  - Numerous interface upgrades
  - New development environment and installation procedures:
    - Creation of a Linux Clustered Environment including a separate Development and Test area
    - Additional storage (100 TB), CPUs available
    - Setup similar to the OPUS environment
    - New software development progression (Dev -> Test -> Ops) used since DR2.5
- Strong presence at AAS; interactions with possible partners in both scientific and broader community
  - Partnerships with Chandra and Spitzer centers
  - Prototypes of enhanced footprint/overlay features using Google Sky, World Wide Telescope, FLEX (Flash-based) software tools

Overview of Data Release 2

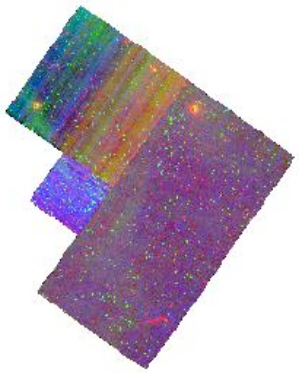
- DR2 took place September 4, 2008
- Major data items:
  - WFPC2 images from CADC
  - Re-processed ACS from EDR
  - ACS DAOPHOT and Source Extractor for all usable ACS images.
- Interface and software improvements
  - Shopping Cart
  - Overlays with SDSS, 2MASS, GSC2 and FIRST
  - Enhanced Plotting Tool
  - User-defined search list
  - Search on prop ID, spectral element, or moving target

DR2 - WFPC2 Details

- WFPC2 Data looks very good
  - Zeropoints understood and match ground based to ~0.03 mag.
  - Photometry is at least good to RMS ~0.05.
    - RMS<0.01 mag for bright objects
    - RMS ~0.05 mag for faint objects
  - WFPC2 data has potential to be as accurate as ACS for relative photometry
- Accuracy needed by some researchers may require that they go back to original chip by chip data
  - Should be a relatively small percentage
  - HLA documentation provides information to allow users to make this decision.



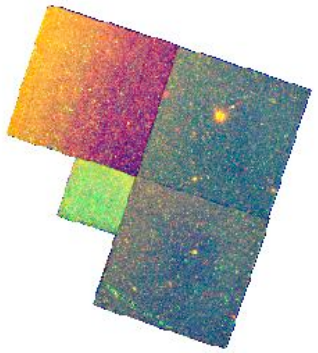
NGC7793 (color)
WFPC2 F656N/F547M PC1
08591_40
Interactive Display
Download Data: FITS-Color (27.6 MB)
Download Source Lists: None
More...



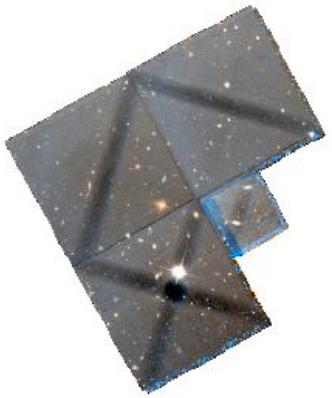
ANY (color)
WFPC2 F814W/F439W/F336W WF2
10246_16
Interactive Display
Download Data: FITS-Color (53.0 MB)
Download Source Lists: None
More...

WFPC2 Images

- In general, the image quality of the ACS and WFPC2 looks similar (e.g., fraction of misalignments, absolute astrometry, etc.)
- The WF4 bias anomaly makes some images look strange
  - Affected images can have bias stripes, appearing as different levels (and thus quite apparent in color composites), and sources fainter than in the other chips
  - The WFPC2 pipeline reprocessing project has specific code for better processing of anomalous bias images; the new data will be ingested in the HLA later this year
- Other known anomalies include time-dependent chip separation, which can make alignment difficult, and the crossed appearance due to earthshine
  - Some of these anomalies will also be improved with the new pipeline, others will not
- A comprehensive ISR describing known WFPC2 image anomalies is linked to our help page



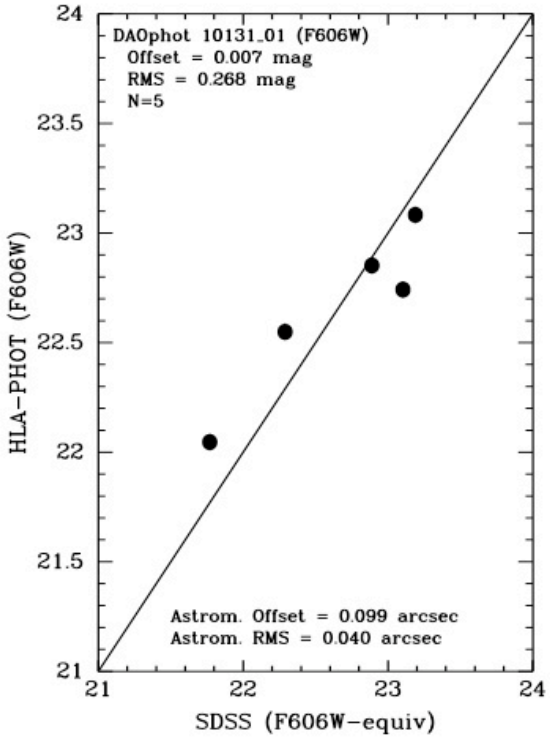
M81-HALO-FIELD (color)
WFPC2 F814W/F555W/F336W WFALL
10136_01
Interactive Display
Download Data: FITS-Color (43.6 MB)
Download Source Lists: None

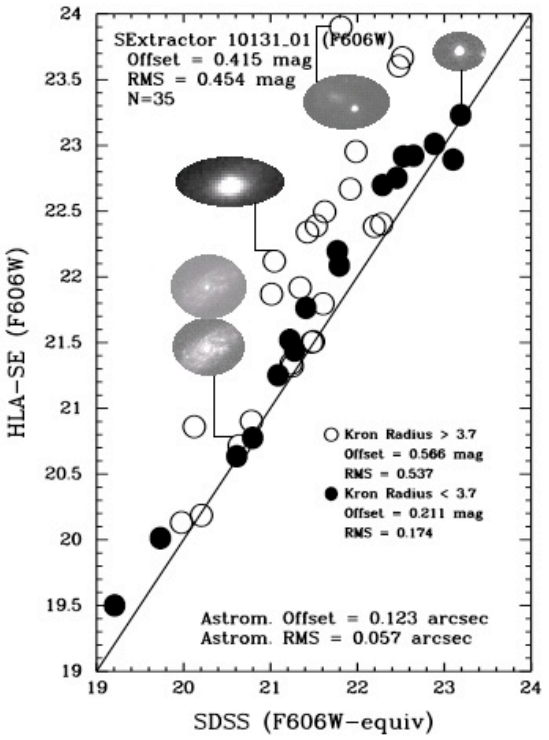


ANY (color)
WFPC2 F814W/F555W WFALL-FIX
10227_02
Interactive Display
Download Data: FITS-Color (50.6 MB)
Download Source Lists: None
More...

ACS Source Lists - Comparison with SDSS

- DAOPHOT zeropoints look ok, although hard to be precise due to large RMS caused by SDSS.
- SExtractor zeropoints also ok. Concentrated objects (i.e., Kron radius < 3.5) match better than extended objects, as expected.
- Probably due to poorer ground based resolution (e.g., ground based cannot separate objects as well as HST)





Overview of Data Release 2.5

- DR2.5 took place on December 22, 2008
- Focus on the addition of Contributed Products
  - User-generated image combinations from the MAST High Level Science Products
  - Most delivered by Treasury and Archival Legacy programs
  - Over 4500 images, including ACS, WFPC2, and NICMOS
  - Functional quality verification; limited review of scientific accuracy
- High-quality but heterogeneous products required several changes behind the scenes
  - Enhancements to footprint, inventory services, default scaling for grayscale and color images
  - Database remains incomplete for some keywords
  - A small fraction of the data excluded from the delivery because of quality issues

DR2.5 – Other Changes

- Some minor changes were made to the HLA interface and procedures at the time of Data Release 2.5
  - Review and enhancements to help information and FAQ to support new products
  - Improved functionality based on user choices
  - Improved cross-functionality of Inventory and Footprint view
  - Small number of problems detected and corrected
  - New, streamlined deployment process to take advantage of Development and Test environments
  - Precomputed multi-resolution FITS data for faster loading, zooming, panning of large-format images

Table of Contributed Products

Project Name	PI	Instrument(s)[1]	Reference Position	Number of Products
ANGST (ACS Nearby Galaxy Survey)	Dalcanton	ACS	Multiple targets	114
COMA (ACS Treasury Survey of Coma Cluster of Galaxies)	Carter	ACS	12:59:49.45 27:55:21.1 1.2d	69
COSMOS (Cosmic Evolution Survey)	Scoville	ACS, WFPC2, NICMOS	10:00:27.85 02:12:03.5 1.0d	999
GEMS (Galaxy Evolution from Morphologies and SEDs)	Rix	ACS	03:32:30 -27:48:20 0.4d	187
GOODS (Great Observatories Origins Deep Survey)	Giavalisco	ACS	North: 12:36:55 62:14:15 0.3d South: 03:32:30 -27:48:20 0.3d	175
Hubble Heritage	Noll	ACS, WFPC2, NICMOS	Multiple targets	91
STAGES (Space Telescope A901/902 Galaxy Evolution Survey)	Gray	ACS	09:56:02.16 -10:05:20.5 0.4d	80
UDF (Ultra Deep Field)	Beckwith, Stiavelli, Thompson	ACS, WFPC2, NICMOS	03:32:29.45 -27:48:18.4 0.1d	20
APPP (Archive Pure Parallels Program)	Casertano	WFPC2	Multiple targets, all sky	2858
SGAL (Spiral Galaxies)	Holwerda	WFPC2	Multiple targets	96

Notes:

1. Only instruments for which HLSP data are available through the HLA interface are listed.

"Best Available" selection

M51 RA = 202.482194 Dec = 47.231509 r = 0.075000 [13:29:55.727 +47:13:53.43]

Instrument	#Footprints
☑ ACS	184
☐ WFPC2-PC	0
☑ WFPC2	79
☐ STIS	0
☐ NICMOS	0
☐ NICGRISM	0
☐ FOS	0
☐ GHRS	0
DSS Image	◉ On ○ Off

Data Product

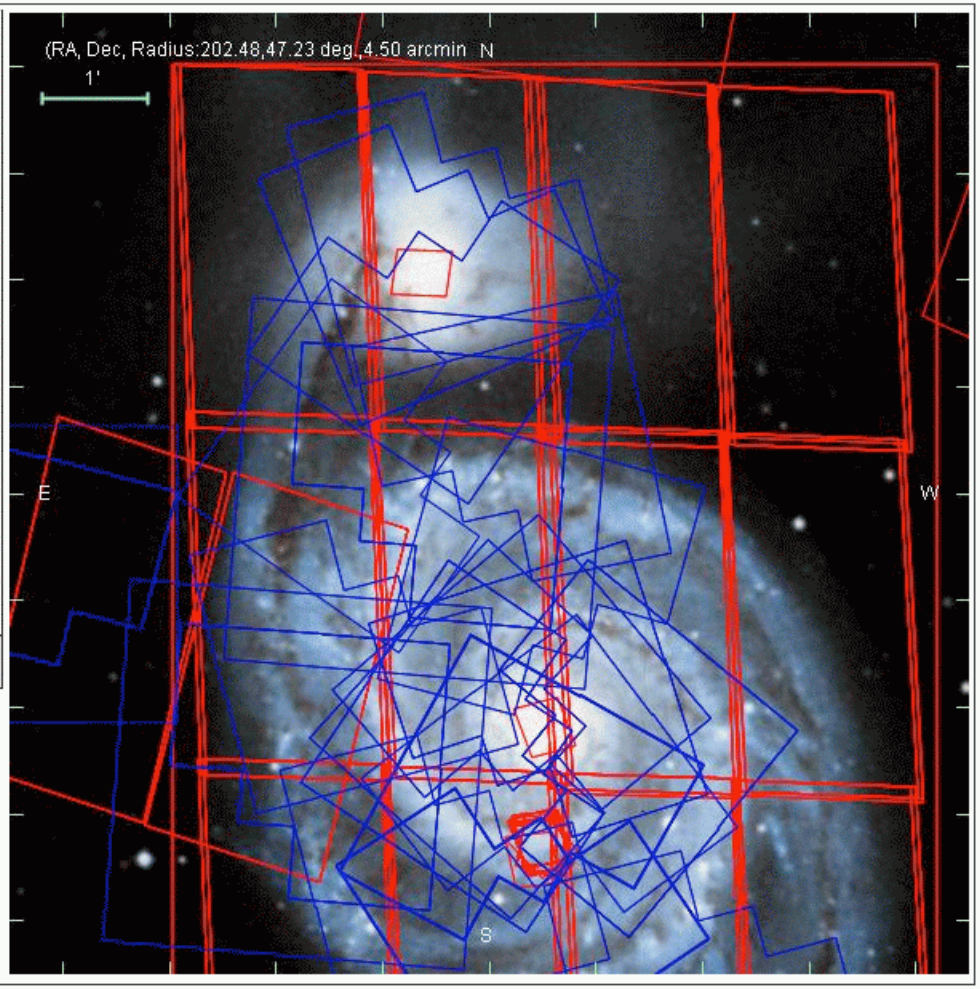
- ○ Exposure(Level 1)
- ○ Combined(Level 2)
- ◉ Best Available
- ○ Contributed HLSP

Submit

To Zoom, go to Advanced Search and enter a smaller value for Radius (smallest value 0.01 degrees)

Click here for NVO STC Web Services

Reset Selection | Save Table



Footprint Science Table (image and table selected fields are sorted to beginning of table)

"Contributed Products" selection

M51 **RA = 202.482194 Dec = 47.231509 r = 0.075000 [13:29:55.727 +47:13:53.43]**

Instrument	#Footprints
☑ ACS	4
☐ WFPC2-PC	0
☑ WFPC2	5
☐ STIS	0
☐ NICMOS	0
☐ NICGRISM	0
☐ FOS	0
☐ GHRS	0
DSS Image	◉ On ○ Off

Data Product

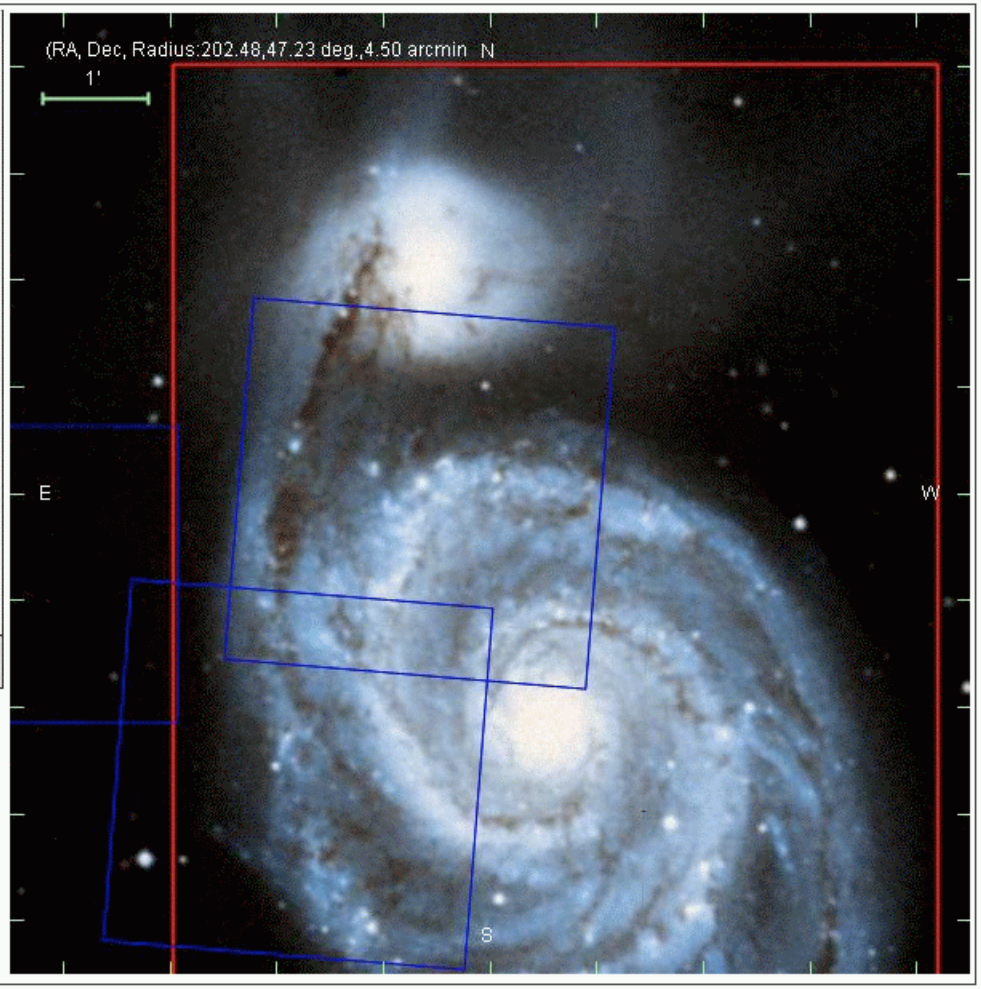
- ○ Exposure(Level 1)
- ○ Combined(Level 2)
- ○ Best Available
- ◉ Contributed HLSP

Submit

To Zoom, go to Advanced Search and enter a smaller value for Radius (smallest value 0.01 degrees)

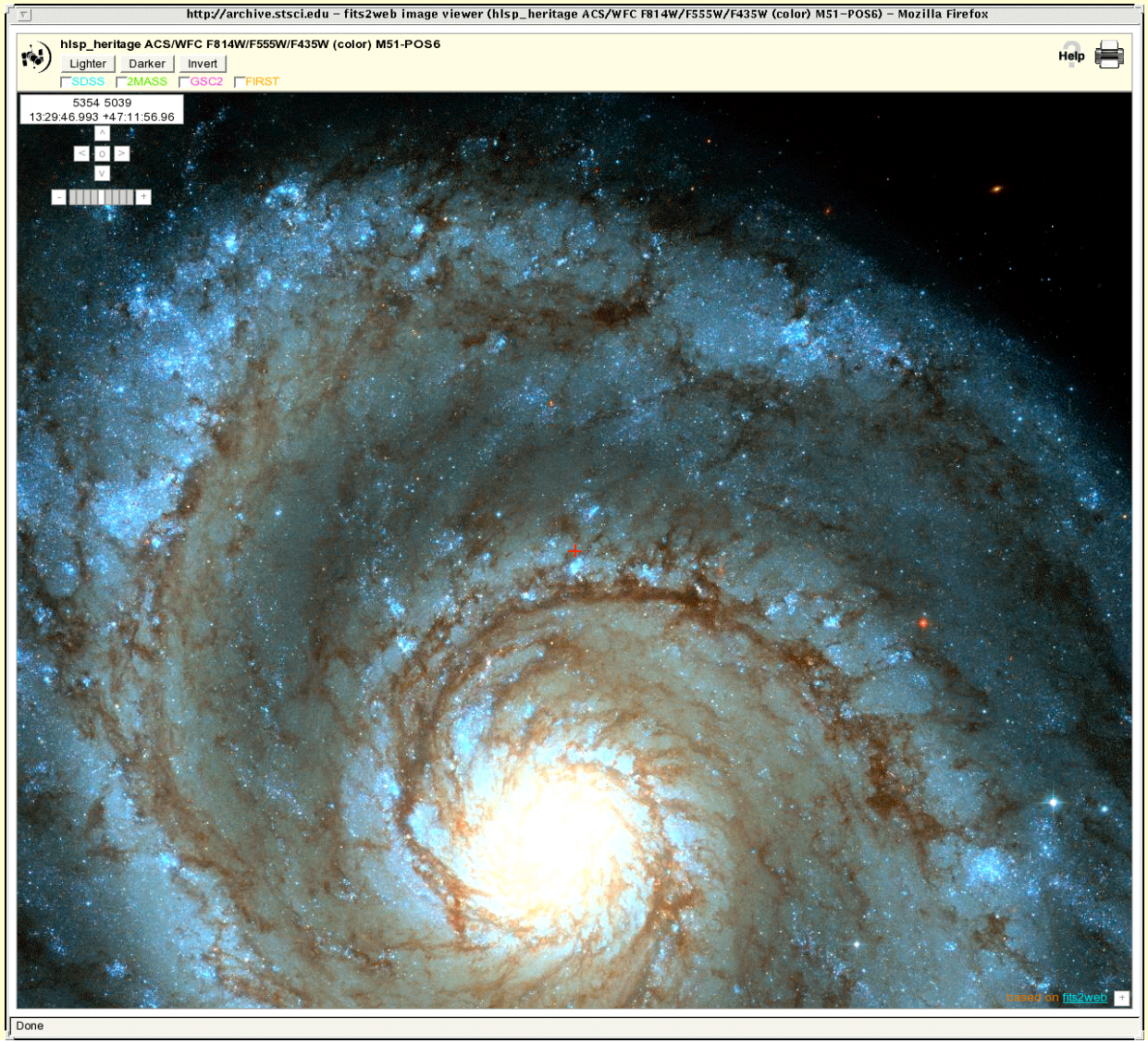
Click here for NVO STC Web Services

Reset Selection | Save Table



Footprint Science Table (image and table selected fields are sorted to beginning of table)

Example: Heritage image of M51



Overview of Data Release 3

- Released May 27, 2009
- New data content:
  - NICMOS reprocessed and visit-combined images
  - ACS Grism data processed by ST-ECF
  - WFPC2 source lists
  - Prototype ACS mosaics (multi-visit combined images)
- Structural and interface enhancements
  - Substantially improved options for sorting and filtering of selected data
  - Full consistency between table views (footprint and inventory display)
  - Enhanced “shopping cart” functionality, integration with DADS requests
  - Complete revision of help and FAQ information
  - GALEX catalog added to overlay options
  - Advanced color controls in Interactive Display

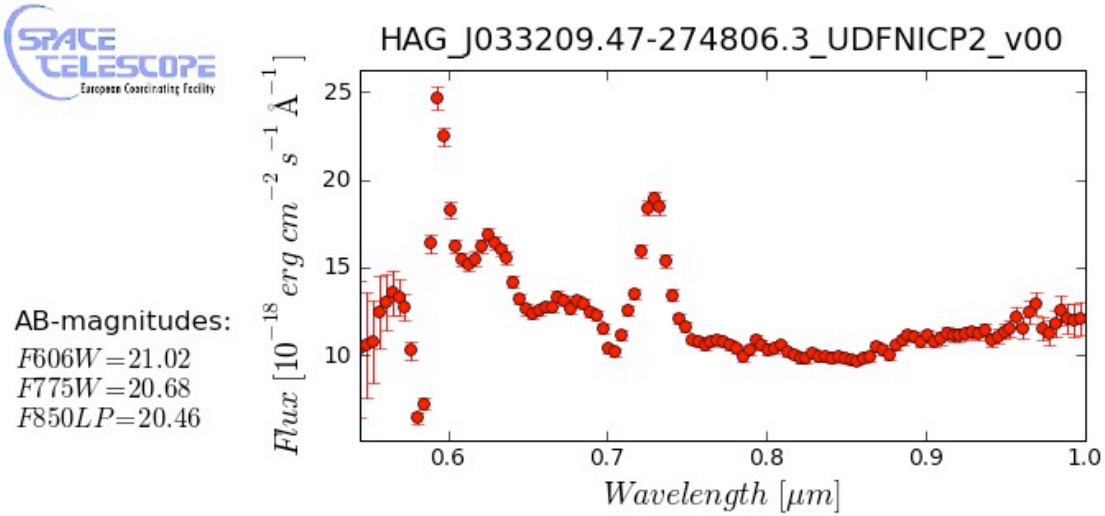
New NICMOS Images

Image of the Carina nebula included in the DR3 data. This color image is produced from F110W and F160W data obtained with NIC3 with a total depth of one orbit.

- With recent pipeline improvements, STSCI is reprocessing all NICMOS data (currently 40% complete)
- Koekemoer, Gaffney developed post-pipeline scripts to obtain Multidrizzle combination of reprocessed data
- All combined images reviewed by NICMOS team in order to validate pipeline improvements
  - Review directly on HLA candidate release images
  - Special web tools provided to facilitate review process
- Results very satisfactory; few images rejected
- DR3 included 25% of all NICMOS images (pre-NCS data)
- Benefits user community as well as NICMOS team (by facilitating quality verification)

ACS Grism data

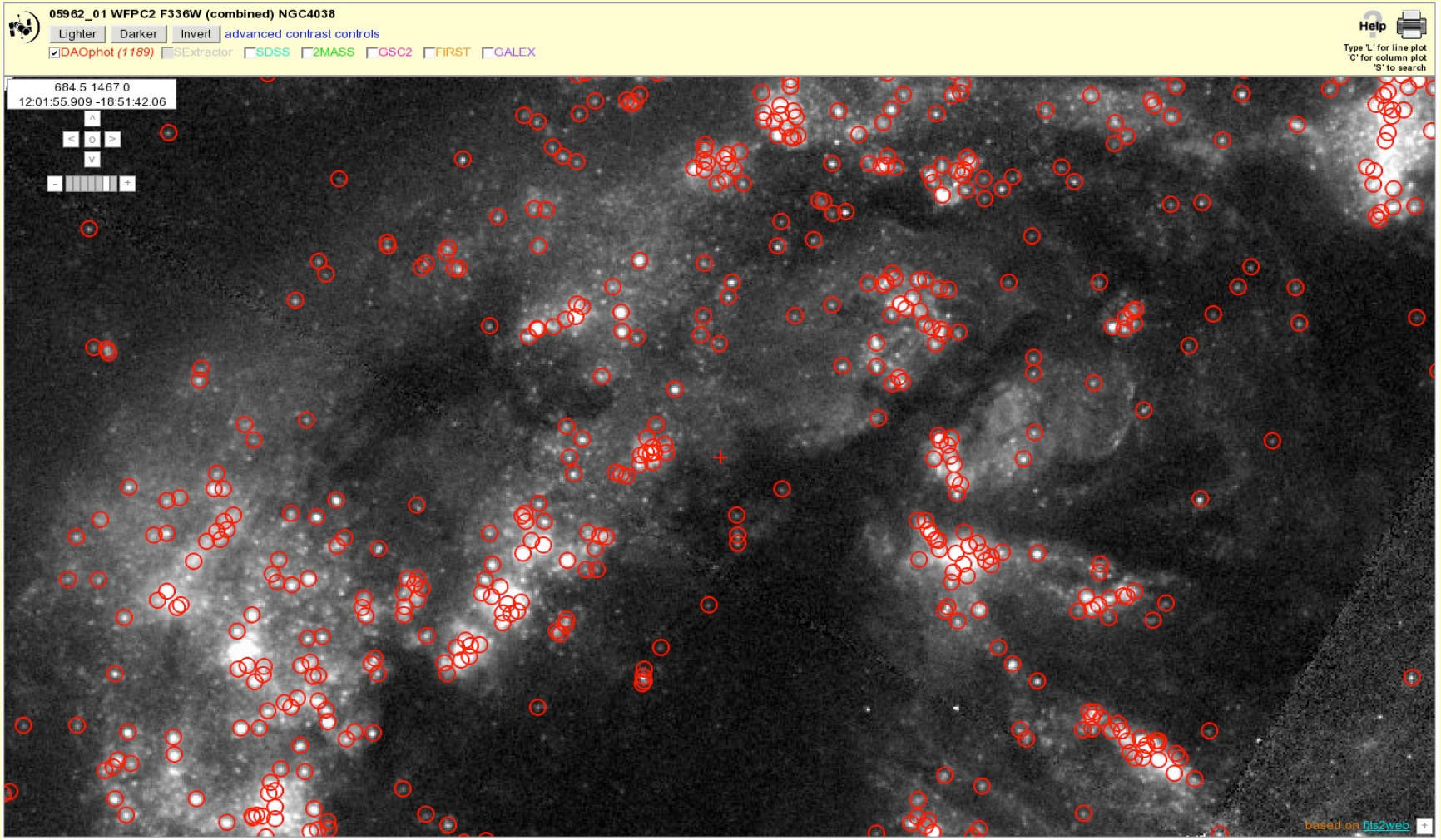
- ST-ECF carried out extraction of 1-dimensional spectra from ACS Grism data
- DR3 included 1235 spectra from the UDF region – estimated about 10% of all GRISM data potentially available
- Data are provided with previews – 1-d plot, direct image, and 2-d spectral image
- Data hosted directly by ST-ECF, but fully integrated in footprint database



Preview of the extracted GRISM spectrum of a z=2.82 BAL QSO included in the DR3 data. The preview includes a plot of the spectrum, a direct image, and a cutout of the two-dimensional GRISM data. A similar preview is available for each extracted target.

WFPC2 Source Lists

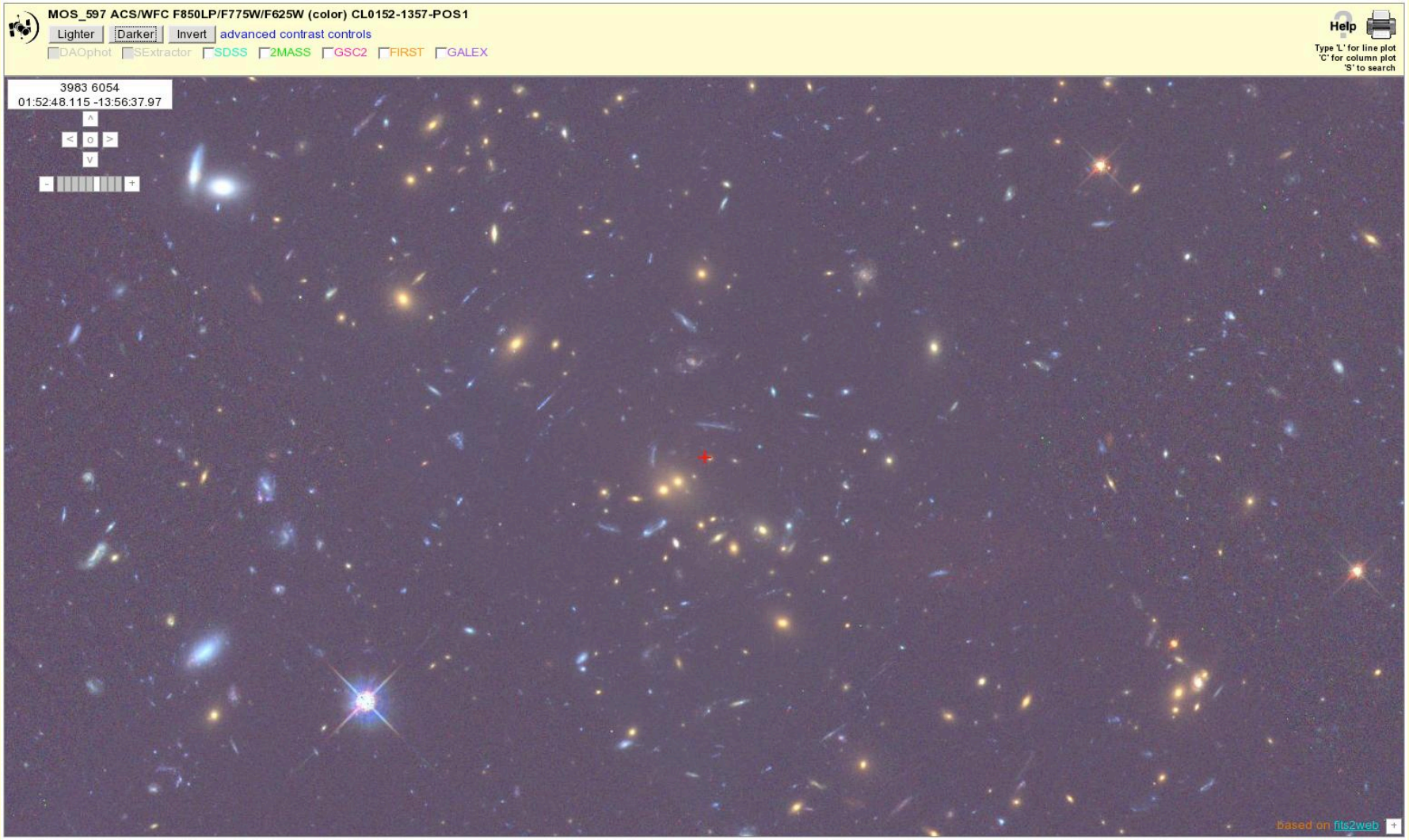
- Same structure as ACS Source Lists
  - Both DAOPhot and Source Extractor lists available
  - Obtained for over 95% of reprocessed WFPC2 data
  - Include CTE correction using Dolphin (2008) prescription
    - CTE correction based on point-source data
    - Extended source correction under development by ACS/WFPC2 team
    - Data used for CTE correction explicitly included in catalog files
- Very good consistency between magnitudes, positions in two sets of sources; match well to independent catalogs in spot checks
- However: Source Extractor lists not deep enough, currently working on improved parameter settings



A region of the WFPC2 image of the Antennae with the an overlay of the DAOPhot sources included in DR3

Prototype mosaics

- Combine overlapping multi-visit data
  - Unlike single-visit data, images cannot be registered from header information
  - New pipeline include astrometry matching step
  - First batch requires at least 4 exposures per filter per visit and narrow (4-month) time window
    - Ensure CR rejection from single-visit information
    - Not sensitive to long term changes in instrument
  - Future developments will relax such restrictions
- Only 8 mosaics available at release
  - Initial quality verification successful encouraging
    - Photometry and pixel-level data consistent with independent combination from ACS Science Team
    - PSF width within 3% of other combinations
    - More quality tests planned in near future with larger data set
- About 100 mosaics processed since release, awaiting full quality verification
  - 90% are expected to be released in current state; 10% have problems that will need reprocessing
  - Enhancements to pipeline may be needed
- We are very interested in feedback on usefulness, formats, quality, ancillary information that may be useful to typical users



Color image of CL0152-1357 generated automatically by the HLA mosaic pipeline. The color image is generated from F625W, F775W, and F850LP, for a total of 57200 s in 8 visits laid out in a 2x2 grid. This is one of 8 mosaics released with DR3; processing of these mosaics is fully automated.

Future plans

- Drivers:
  - Community input (User Survey, MUG, help requests, user comments)
  - HLA Board review
  - Long-term plan to enhance quality, consistency, scope of products
  - Fix known issues
- Tools:
  - Issue-tracking system
    - Classify issues as fixes, enhancements, new tasks
    - Identify severity/priority, time scale, person responsible
    - Allows comments and progress tracking
    - Various types of reports available
    - Lightweigth system, very little overhead
  - Wiki pages for more complex planning
  - Basecamp for file sharing, meeting management
  - Frequent review of priorities, status
- High-level items recommended by MUG:
  - Enhance visibility of HLA within the community
  - Improve bandwidth at STScI
  - Identify underutilized data – in general, track use of archival data and extract appropriate information
  - Improve treatment of spectral data

Future plans (cont.)

- Short term (1-2 months): complete the work started for DR3
  - Improve WFPC2 SExtractor source lists (ongoing)
  - Add more sample mosaics (already up to 100, pending quality verification)
  - Include more NICMOS data as reprocessing continues (ongoing)
  - Resolve known issues
- Medium term (3-6 months): Improve existing data products, prepare for new instruments
  - Revisit ancillary data, keyword information; modify products when necessary for better fidelity
  - Extend absolute astrometry enhancement to all HST data
    - Expected worst-case accuracy 0.4”
    - Typical accuracy 0.1-0.2”
  - Work with Instrument Teams to understand best products for WFC3, COS, possibly STIS
  - Define content, timetable for DR4; include products for new instruments as soon as practical
  - Incorporate reprocessed WFPC2 data with new pipeline
  - Move towards a continuous data flow model to release products in small batches
  - Improve mosaic pipeline to relax current restrictions
    - Carefully weigh resources needed for more general mosaic generation vs. other developments on the basis of community demand
  - Continue work with Chandra, Spitzer to integrate our footprint data
  - Support footprint service for APT
  - Enhance our quality testing
    - Systematic intercomparison of repeated images and cross-instrument checks
    - Routine comparison with existing ground-based photometry
    - Production of additional diagnostics on alignment, photometry during pipeline processing
    - Statistics of source distribution and properties to identify anomalous images
- Long term (1 year+):
  - Evaluate new software tools for enhanced interface
    - Prototypes allow concurrent display of multiple image layers, contextual data through browser
    - Stability, availability, and resource demands of software will be considered
  - Provide enhanced products for spectral data