

# File Format: BGP (Affymetrix Power Tools v1.6.1)

The BGP (background probe file) lists which probes to use for background correction. Currently this file is used with the PM-GCBG pm adjustment method implemented in apt-probeset-summarize.

The BGP file is based on version 1 of the **TSV file format**.

## Specs & Features

- Required headers
  - chip\_type: indicates the chip type (as stored in the CEL file) which the clf file describes. Multiple chip\_type headers may be present indicating the clf file can be used with any one of the listed chip types.
  - lib\_set\_name: indicates the name of a collection of related library files for a given chip. For example, PGF and CLF files intended to be used together should have the same lib\_set\_name. Only a single lib\_set\_name is allowed.
  - lib\_set\_version: indicates the version of a collection of related library files for a given chip. For example, PGF and CLF files intended to be used together should have the same lib\_set\_version. Only a single lib\_set\_version is allowed.
- Optional headers
  - create\_date: timestamp for when the file was created
  - guid: a unique identifier for the file
  - other headers may be present
- Required columns
  - probe\_id: the ID of the probe to use for background correction
- Optional columns
  - probeset\_id: The ID of the probeset in the PGF file to which the probe belongs
  - probeset\_type: See type field at probeset level in PGF documentation
  - atom\_id: The ID of the atom to which the probe belongs to
  - probe\_type: See type field at the probe level in PGF documentation
  - gc\_count: The numbers of G's and C's in the probe
  - probe\_length: The length of the probe in base pairs
  - probe\_sequence: See the probe\_sequence info in the PGF documentation
  - x: The column coordinate of the probe in the CEL file
  - y: The row coordinate of the probe in the CEL file
  - other columns may be present

## Parsing and Writing

The official C++ parser used by affy can be found in APT under `sdk/file/TsvFile/BgpFile.h`. When possible, parsing and writing of BGP files should be done using this code.

## Notes

The optional fields listed above are frequently present. This is in part due to the fact that `apt-dump-pgf` is used to generate the BGP files.

## Example 1

```
##chip_type=HuEx-1_0-st-v2
##chip_type=HuEx-1_0-st-v1
##chip_type=HuEx-1_0-st-ta1
##lib_set_name=HuEx-1_0-st
##lib_set_version=r2
##create_date=Tue Sep 19 15:18:05 PDT 2006
##guid=0000008635-1158704285-2136082293-1007618462-0634896595
probeset_id      probeset_type  atom_id probe_id
probe_type      gc_count      probe_length  probe_sequence
x              y
4057134 control->bgp->antigenomic      70013  78973  pm:st
4              25          TATTAAAATTAGAGCTATTAACAAT    2172   30
4057134 control->bgp->antigenomic      70013  125043 pm:st
4              25          AACTTATCTAAATGCTATTA AAAAT    2162   48
4057134 control->bgp->antigenomic      70013  148149 pm:st
4              25          TGATATAATAAGCTAATTTGTTTAA    2228   57
4057134 control->bgp->antigenomic      70013  152379 pm:st
4              25          ATTATAATAGTTAGCAACTTAATAT    1338   59
4057134 control->bgp->antigenomic      70013  178557 pm:st
4              25          CATTTATTAGTATTAAAGCAAATAT    1916   69
4057134 control->bgp->antigenomic      70013  178785 pm:st
4              25          AATTATTAACAAGCTTATATATC      2144   69
```

## Related Pages

- [TsvFile Design Notes](#)
- [TSV File Format](#)

