

Establishing new diffuse interstellar band correlations to identify common carriers

Fraser M. Smith,¹ Tina A. Harriott,^{1★} Daniel Majaess,^{2★} Lou Massa³ and Chérif F. Matta^{4,5,6★}

¹Department of Mathematics and Statistics, Mount Saint Vincent University, Halifax, Nova Scotia, B3M 2J6 Canada

²Department of Chemistry and Physics, Mount Saint Vincent University, Halifax, Nova Scotia, B3M 2J6 Canada

³Hunter College & the PhD Program of the Graduate Center, City University of New York, NY 10065, New York, USA

⁴Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4J3 Canada

⁵Department of Chemistry, Saint Mary's University, Halifax, Nova Scotia, B3H 3C3 Canada

⁶Département de Chimie, Université Laval, Québec, Québec, G1V 0A6 Canada

Accepted 2021 August 16. Received 2021 August 5; in original form 2021 May 10

ABSTRACT

Observations from the *Apache Point Observatory Catalog of Optical Diffuse Interstellar Bands* (DIBs) were analysed to establish highly correlated pairs in terms of their equivalent widths (EWs) ($r > 0.95$), which importantly facilitate the identification of common carriers. A total of 154 846 possible DIB pairs were originally examined, yet only those with a sufficient number of sightlines ($n > 9$) that included EW uncertainties were subsequently investigated. The highest correlations for the resulting 56 893 DIB pairs are 6284.05–6203.58 Å ($r = 0.990 \pm 0.001$), 6203.58–5780.64 Å ($r = 0.986 \pm 0.001$), 6993.12–6269.89 Å ($r = 0.984 \pm 0.001$), 6843.76–6792.51 Å ($r = 0.984 \pm 0.005$), 6203.58–5487.64 Å ($r = 0.983 \pm 0.002$), and 5061.50–4969.12 Å ($r = 0.983 \pm 0.009$). The bands 5363.77, 5780.64, 6203.58, and 6284.05 Å appear most frequently. Novel relations linked to those DIBs and others warrant further research, in particular those pairs that involve one or both DIBs with low EWs (e.g. 5609.82, 6269.89, 6993.12, and 7224.16 Å). Numerous DIBs correlated with the prominent 4429.33 Å band were also discovered. The intriguing proposal of anionic hydrogen clusters as possible DIB carriers is also discussed.

Key words: ISM: clouds – ISM: lines and bands – ISM: molecules.

1 INTRODUCTION

There is currently no consensus regarding the source(s) behind the diffuse interstellar bands (DIBs). More than 559 DIBs have been identified to date (Fan et al. 2019). Examining the relations between observed DIBs provides critical insight to benchmark proposed sources, namely via a comparison to theoretical (quantum chemical) or experimentally derived electronic spectra.

DIBs exhibiting high correlations between their intensities, equivalent widths (EWs), and spectral profiles may constitute a DIB family. Krelowski & Walker (1987) proposed that the DIBs centered at 4430, 5780, 5797, 5850, 6180, 6196, 6203, 6269, 6284, 6376, 6379, and 6614 Å may belong to three groups based on relative intensities. Correlations between DIB EWs led Cami et al. (1997) to identify two families. Moutou et al. (1999) identified correlations between the central depths of certain DIB pairs, the highest being 6614–6196, 6614–5780, 6196–5780, 5850–5797, 6196–5797, 6379–5797, 6614–5797, and 6379–6196 Å. That work lent support to a family proposed by Cami et al. (1997), on the basis of the correlations between 5797, 6379, and 6614 Å (Moutou et al. 1999). Friedman et al. (2011) discovered that the highest correlation ($r = 0.99$) is between 6196.0 and 6613.6 Å (see also McCall et al. 2010). Xiang, Liu & Yang (2012) cited strong correlations ($r \geq 0.95$) between 11

DIB pairs, with certain bands being classified into broad or narrow (full width at half-maximum) families.

More recently, Bondar (2020) identified high correlations ($0.968 \leq r \leq 0.988$) between the EWs of 5545, 6113, 6196, 6445, and 6614 Å. Bondar (2020) argued that the wavenumber spacing implied an association between those DIBs and polycyclic aromatic hydrocarbons. Fullerenes are also considered as potential DIB carriers (i.e. C_{60}^+), however, Galazutdinov & Krelowski (2017) report varying DIB strength ratios associated with the putatively associated pairs 9632–9577 and 9577–9366 Å. That is one of the contested issues associated with C_{60}^+ , and the reader is referred to Linnartz et al. (2020) for an alternate viewpoint. Anionic hydrogen clusters have also been proposed as a source of DIBs (Huang et al. 2019), and the abundance of hydrogen and free electrons in the interstellar medium makes them attractive candidates.

In this study, observations from the *Apache Point Observatory Catalog of Optical Diffuse Interstellar Bands* (Fan et al. 2019) were analysed to establish highly correlated pairs. The principal aim being to provide observational constraints to potentially identify DIB families and benchmark proposed carriers.

2 OBSERVATIONS AND METHODS

The DIB database recently published by Fan et al. (2019) was used. The database features 25 stars observed through moderate to high extinction ($0.31 < E(B - V) < 3.31$), and spanning diverse interstellar

* E-mail: Tina.Harriott@msvu.ca (TAH); Daniel.Majaess@msvu.ca (DM); Cherif.Matta@msvu.ca (CFM)

Table 1. Stars found in the *Apache Point Observatory Catalog of Optical Diffuse Interstellar Bands* (an abridged version of table 1 from Fan et al. 2019).

Star	SpT	$E(B - V)$	f_{H2}
HD 20041	A0Ia	0.72	0.42
Cernis 52	A3V	0.90	>0.78
HD 23180	B1III + B2V	0.31	0.55
HD 281159	B5V	0.85	0.50
HD 23512	A0V	0.36	0.62
HD 24534	O9.5pe	0.59	0.76
HD 24912	O7e	0.33	0.38
HD 28482	B8III	0.48	0.66
HD 37061	B1V	0.52	0.02
HD 37903	B1.5V	0.35	0.53
HD 43384	B3Ib	0.58	0.44
HD 147084	A5II	0.73	0.59
HD 147889	B2V	1.07	0.45
HD 148579	B9V	0.34	0.45
HD 166734	O8e	1.39	0.39
HD 168625	B8Ia	1.48	0.33
HD 175156	B5II	0.31	0.31
HD 183143	B7Iae	1.27	0.31
HD 190603	B1.5Iae	0.72	0.16
HD 194279	B2Iae	1.20	0.30
VI Cyg 5	O7f	1.99	0.47
VI Cyg 12	B5Ie	3.31	>0.48
HD 204827	O9.5V + B0.5III	1.11	0.67
HD 206267	O6f	0.53	0.42
HD 223385	A3Iae	0.67	0.12

conditions and spectral types (O to A). Table 1 contains the stellar designations and spectral types for the stars analysed (see also Fan et al. 2019).

A total of 559 DIBs were identified by Fan et al. (2019) between 4000 and 9000 Å, however, this number includes two bands that were listed but not measured and consequently not considered in this work. The data were collected from the Apache Point Observatory using an ARC echelle spectrograph on the 3.5 m telescope.

Correlations between bands were determined using the standard Pearson coefficient (r):

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (1)$$

The correlations were computed using SciPy’s `pearsonr` function (Virtanen et al. 2020), and plots were created using Matplotlib (Hunter 2007). A total of 154 846 = [(557 × 557) − 557] / 2 correlations according to equation (1) between DIBs were considered. DIB pairs possessing EW uncertainties and sufficient data ($n > 9$) were examined. That yielded 56 893 DIB pairs, whereby a subset of 231 pairs exhibit $r > 0.95$. To mitigate artefacts introduced by low-sampling, a higher threshold ($n \geq 15$) was adopted for DIB pairs conveyed in Tables 2 and 3, which are assigned a higher degree of confidence. Correlations found in Tables 4 and 5 are deemed tentative owing to reduced statistics ($15 > n > 9$). Independent observations are needed to verify the suggested pairs. Fig. 1 displays a subset of highly correlated and understudied DIBs.

It is important to consider uncertainties, as a correlation may be overemphasized if its unreliability is overlooked. Uncertainties in the correlation coefficients and slopes were obtained through Monte Carlo simulations; 1000 datasets were randomly generated for each DIB pair using the EW uncertainties. The points were randomly distributed (uniformly distributed, on average) within the

uncertainties. The correlation coefficient and slope of the linear regression were evaluated for each dataset, and the standard deviations were subsequently adopted as the uncertainties. The largest uncertainties tend to arise in pairs containing weak DIBs.

The analysis relies on formal underestimated EW uncertainties. Numerous systematic uncertainties can arise that affect EW measurements such as telluric line contamination, blending, continuum placement, and stellar properties (Bondar 2020; Friedman et al. 2011). Systematic uncertainties will be considered in a subsequent study.

3 DIB CORRELATIONS

The maximally correlated pairs of the 557 DIBs observed are discussed below, and indeed, numerous pairs linked to various lines were possibly identified for the first time (e.g. 5236.27, 5363.77, 5609.82, 6269.89, 6993.12, and 7224.16 Å).

The highest correlation derived is between 6284.05 and 6203.58 Å ($r = 0.990 \pm 0.001$). The correlation between 5779.59 and 5780.64 Å supersedes the aforementioned result, however, there are concerns regarding the deblending of nearly coincident lines. The DIB 5779.59 Å is the broad feature measured over the entire 5780 Å region, whereas 5780.64 Å is sharper and deeper (Fan et al. 2019). Fig. 4 in Galazutdinov et al. (2020) provides a visual profile of the two DIBs. The DIB 5779.59 Å is omitted from the present work and will be considered in a subsequent analysis.

The average correlation offset between the 10 DIB pairs in common with Bondar (2020) is marginal (0.059 ± 0.041), which points to a general agreement. For example, the correlation determined here between 6113.22 and 6613.74 Å is $r = 0.960 \pm 0.003$, which is comparable to the Bondar (2020) result ($r = 0.982$).

Several DIBs were common among the most highly correlated pairs. Tables 3 and 5 convey the appearance rate of such DIBs. The DIB that appears most frequently is 5363.77 Å, a band that is scarcely discussed. Of the 15 highest correlations found in the database, 5780.64 Å appears most often (with four pairs among the 15 highest correlations) followed by 6203.58 and 5363.77 Å (each with three pairs). Correlations involving the DIB 5363.77 Å are tied to lower sample sizes than the $n = 15$ threshold, relegating such pairs to a lower degree of confidence. A larger database of sightlines is required to confirm the DIB pairs with smaller sample sizes.

The DIB pairs 6993.12–6269.89, 6843.76–6792.51, 5061.50–4969.12, 6498.00–6245.14, and 6330.03–6245.14 Å are deserving of further research since the correlations exhibited by these pairs are among the highest compared to all possible pairs of the DIBs examined in this study. In addition, these DIBs in particular appear to be overlooked in the literature.

It is also of interest that the strongest known DIB, 4429.33 Å, was well-correlated with five DIBs. Specifically, high correlations were found between 4429.33 and 6624.90 Å as well as 4429.33 and 5236.27 Å, though in this instance a lower threshold than $r > 0.95$ was adopted given the prominence of the former band. Fig. 2 shows the plots for these two pairs. A firm conclusion regarding this set of DIBs is not justified at this time owing to a lack of observations, especially for 6624.90 Å. Table 6 lists the correlations between each possible pair of bands in this set.

DIBs that appear frequently among the most highly correlated pairs (i.e. DIBs with numerous entries in Tables 2 and/or 4), and that also exhibit high correlations between one another, were studied further. Three sets of well-correlated DIBs are identified; one set notably contains the weak DIB 5363.77 Å (the most common DIB among highly correlated pairs), another set is centred around two

Table 2. The highest correlated DIB pairs identified featuring $n \geq 15$.

λ_1 (Å)	λ_2 (Å)	$r \pm \sigma$	n	λ_1 (Å)	λ_2 (Å)	$r \pm \sigma$	n	λ_1 (Å)	λ_2 (Å)	$r \pm \sigma$	n
6284.05	6203.58	0.990 ± 0.001	25	6108.06	5494.10	0.963 ± 0.003	17	5780.64	5609.82	0.955 ± 0.004	19
6203.58	5780.64	0.986 ± 0.001	25	7559.43	6843.76	0.963 ± 0.007	17	5797.18	5494.10	0.955 ± 0.002	20
6993.12	6269.89	0.984 ± 0.001	22	7562.16	5609.82	0.963 ± 0.005	18	7559.43	6494.13	0.955 ± 0.011	15
6203.58	5487.64	0.983 ± 0.002	20	5418.87	4984.78	0.963 ± 0.004	17	6211.69	6195.99	0.955 ± 0.003	20
6613.74	6195.99	0.980 ± 0.001	24	6009.60	5795.21	0.962 ± 0.005	18	6795.26	6113.22	0.955 ± 0.005	17
5780.64	5705.12	0.980 ± 0.001	20	7334.53	6492.01	0.962 ± 0.004	15	5769.09	5176.00	0.955 ± 0.006	15
6284.05	5780.64	0.978 ± 0.001	25	7224.16	5609.82	0.962 ± 0.003	17	7559.43	6770.17	0.955 ± 0.006	17
7562.16	6203.58	0.977 ± 0.001	21	6269.89	5780.64	0.961 ± 0.001	24	6139.95	6113.22	0.955 ± 0.005	18
6284.05	5487.64	0.977 ± 0.003	20	6116.80	5545.08	0.961 ± 0.005	18	6439.51	5849.82	0.954 ± 0.002	24
7224.16	6284.05	0.977 ± 0.001	21	5418.87	5176.00	0.961 ± 0.005	16	6439.51	5494.10	0.954 ± 0.004	20
7832.88	6803.35	0.976 ± 0.004	15	7832.88	6993.12	0.961 ± 0.003	18	6324.91	5487.64	0.954 ± 0.004	18
5780.64	5487.64	0.976 ± 0.003	20	7367.08	6185.79	0.960 ± 0.005	19	6211.69	5545.08	0.954 ± 0.003	17
6284.05	5609.82	0.975 ± 0.003	19	5418.87	4963.92	0.960 ± 0.003	20	7367.08	6613.74	0.954 ± 0.002	22
6269.89	6195.99	0.975 ± 0.001	24	7061.09	6843.76	0.960 ± 0.004	17	6919.26	6284.05	0.954 ± 0.006	18
6203.58	5609.82	0.975 ± 0.003	19	6622.84	6185.79	0.960 ± 0.006	21	6353.31	5487.64	0.954 ± 0.006	18
6993.12	5404.58	0.974 ± 0.005	16	6195.99	5236.27	0.960 ± 0.004	15	6445.30	6234.01	0.954 ± 0.005	19
6376.14	6367.30	0.974 ± 0.002	21	6613.74	6113.22	0.960 ± 0.003	20	7832.88	6498.00	0.954 ± 0.006	16
7224.16	6203.58	0.974 ± 0.001	21	6993.12	5780.64	0.959 ± 0.001	23	8026.23	7562.16	0.954 ± 0.003	17
6195.99	5780.64	0.974 ± 0.001	24	6367.30	6113.22	0.959 ± 0.004	19	6919.26	6011.64	0.954 ± 0.005	15
6211.69	6113.22	0.974 ± 0.002	18	5705.12	5487.64	0.959 ± 0.004	18	6194.73	6113.22	0.953 ± 0.005	17
6203.58	5705.12	0.972 ± 0.002	20	6269.89	5508.33	0.959 ± 0.002	15	6377.07	6376.14	0.953 ± 0.003	22
6993.12	6195.99	0.972 ± 0.001	22	6011.64	6009.60	0.959 ± 0.006	20	7334.53	7224.16	0.953 ± 0.002	16
6613.74	6269.89	0.972 ± 0.001	24	6699.28	6367.30	0.959 ± 0.003	21	6811.27	6801.47	0.953 ± 0.006	16
5797.18	5545.08	0.972 ± 0.003	21	6449.27	5494.10	0.959 ± 0.005	19	6993.12	6843.76	0.953 ± 0.004	18
6284.05	5705.12	0.972 ± 0.002	20	7367.08	6702.07	0.958 ± 0.004	19	6672.23	6660.67	0.953 ± 0.004	15
6439.51	5545.08	0.971 ± 0.002	21	6520.74	6211.69	0.958 ± 0.005	19	7062.69	5925.91	0.953 ± 0.008	15
5609.82	5487.64	0.970 ± 0.004	18	6195.99	5900.59	0.958 ± 0.009	17	7224.16	6011.64	0.953 ± 0.004	19
6613.74	6367.30	0.969 ± 0.003	22	7224.16	6740.97	0.958 ± 0.005	16	8026.23	6811.27	0.952 ± 0.004	16
7224.16	6919.26	0.968 ± 0.003	17	5780.64	5236.27	0.958 ± 0.004	15	6801.47	6203.58	0.952 ± 0.005	17
7562.16	6284.05	0.967 ± 0.001	21	4984.78	4963.92	0.957 ± 0.004	19	6993.12	6613.74	0.952 ± 0.001	23
6613.74	6520.74	0.967 ± 0.003	22	6919.26	6203.58	0.957 ± 0.006	18	6843.76	6770.17	0.952 ± 0.005	17
6795.26	5545.08	0.967 ± 0.005	17	6269.89	6065.32	0.957 ± 0.004	21	7559.43	6803.35	0.952 ± 0.009	15
7061.09	6770.17	0.966 ± 0.003	18	7367.08	6367.30	0.957 ± 0.003	20	6702.07	6689.35	0.952 ± 0.006	18
7224.16	6353.31	0.966 ± 0.004	18	5849.82	5797.18	0.957 ± 0.001	25	6185.79	6113.22	0.952 ± 0.004	18
6439.51	5797.18	0.966 ± 0.002	24	6520.74	6269.89	0.957 ± 0.003	22	6520.74	6113.22	0.952 ± 0.004	19
6520.74	6195.99	0.966 ± 0.003	22	6622.84	6113.22	0.957 ± 0.005	19	6268.59	5705.12	0.951 ± 0.004	18
6765.29	5766.16	0.966 ± 0.006	15	5705.12	5609.82	0.956 ± 0.005	16	6919.26	6353.31	0.951 ± 0.008	16
6379.25	6089.85	0.965 ± 0.003	23	6993.12	6203.58	0.956 ± 0.001	23	6195.99	5487.64	0.951 ± 0.004	20
6801.47	5609.82	0.965 ± 0.005	16	4963.92	4726.98	0.956 ± 0.007	24	8026.23	5609.82	0.951 ± 0.006	18
6203.58	6195.99	0.964 ± 0.001	24	6689.35	6367.30	0.956 ± 0.007	18	7224.16	5780.64	0.951 ± 0.001	21
6770.17	6494.13	0.964 ± 0.008	15	7357.60	6597.34	0.956 ± 0.003	18	6449.27	5797.18	0.951 ± 0.004	23
6195.99	6108.06	0.964 ± 0.003	20	6108.06	5797.18	0.955 ± 0.002	21	7061.09	6978.47	0.950 ± 0.005	16
6269.89	5404.58	0.964 ± 0.005	17	6113.22	5797.18	0.955 ± 0.002	20	6919.26	6801.47	0.950 ± 0.005	15
7224.16	6597.34	0.964 ± 0.003	20	6108.06	5925.91	0.955 ± 0.008	18	6770.17	6597.34	0.950 ± 0.004	18
5849.82	5545.08	0.964 ± 0.002	21	7224.16	5487.64	0.955 ± 0.003	19
7367.08	6689.35	0.964 ± 0.003	17	7562.16	5487.64	0.955 ± 0.004	19

weak DIBs 5609.82 and 7224.16 Å, and the final set is based around correlations with 6269.89 and 6993.12 Å. The Pearson correlation coefficients (r) between each pair of the DIBs are listed in matrix format in Tables 7–9. Each table relays the uncertainties of each r -value. Accompanying figures can be found in the Appendix (Figs A1–A3).

The three sets of well-correlated DIBs are as follows: 5363.77, 5487.64, 5705.12, 5780.64, 6203.58, and 6284.05 Å form one set, 5487.64, 5609.82, 5780.64, 6203.58, 6284.05, and 7224.16 Å form another set, and 5780.64, 6195.99, 6269.89, 6613.74, and 6993.12 Å form the last set. The first set was obtained by considering all DIBs that appear more than 5 times in Table 2 or Table 4, and removing the DIB with the lowest correlations between each of these DIBs, repeating until the lowest correlation in the set was $r \geq 0.95$. The number of appearances of selected DIBs in Tables 2 and 4 are listed in

Tables 3 and 5, respectively. The DIB 5487.64 Å was later appended to this set, based on the high correlations between 5487.64 Å and every other DIB in the set (as indicated by the correlations found in the following set). The second set was identified by considering two bands, 5609.82 and 7224.16 Å, and checking any DIB that was well-correlated with either of the two bands, including the DIB if it was well-correlated with every DIB in the expanding set. These two particular bands were chosen because they are understudied, and since they are both well-correlated with the previously studied DIBs 5487.64, 5780.64, 6203.58, and 6284.05 Å. The final set was identified by considering all bands well-correlated ($r > 0.95$) with either of the DIBs 6269.89 or 6993.12 Å and performing a similar procedure as for the previous set. These two bands were chosen owing to their high correlation with 6245.14 Å, which does not share many

Table 3. The most frequently identified DIB bands among highly correlated pairs featuring $n \geq 15$.

DIB (Å)	Appearance/count
6195.99	11
6203.58	11
7224.16	11
5487.64	10
5780.64	10
6113.22	10
5609.82	9
6269.89	8
6284.05	8
6993.12	8
5797.18	7
6613.74	7
5545.08	6
5705.12	6
6367.30	6
6919.26	6

of the same high correlations with certain DIBs as those considered in the previous two sets, and since these two weak bands share a high correlation with many of the same DIBs. Although both 6269.89 and 6993.14 Å have high correlations with 6245.14 Å, 6245.14 Å is absent from the final set of DIBs due to a slightly lower correlation with other DIBs in the set (0.924 ± 0.005 and 0.938 ± 0.004).

Table 4. The highest correlated DIB pairs identified featuring $15 > n > 9$.

λ_1 (Å)	λ_2 (Å)	$r \pm \sigma$	n	λ_1 (Å)	λ_2 (Å)	$r \pm \sigma$	n	λ_1 (Å)	λ_2 (Å)	$r \pm \sigma$	n
6843.76	6792.51	0.984 ± 0.005	13	6245.14	6065.32	0.966 ± 0.004	11	7405.73	6594.30	0.956 ± 0.015	11
5061.50	4969.12	0.983 ± 0.009	12	6594.30	5705.12	0.966 ± 0.009	12	6803.35	6494.13	0.956 ± 0.014	14
6498.00	6245.14	0.983 ± 0.005	11	6788.83	6220.95	0.965 ± 0.013	11	7366.05	5363.77	0.956 ± 0.011	10
6489.42	5719.57	0.982 ± 0.006	13	5711.46	4501.51	0.965 ± 0.003	10	5705.12	5363.77	0.956 ± 0.007	11
6330.03	6245.14	0.982 ± 0.005	11	6203.58	5363.77	0.964 ± 0.005	13	5925.91	5363.77	0.955 ± 0.010	12
5487.64	5363.77	0.981 ± 0.005	12	6089.85	5821.15	0.964 ± 0.008	12	6195.99	5363.77	0.955 ± 0.005	13
5780.64	5363.77	0.980 ± 0.003	13	7061.09	6498.00	0.963 ± 0.007	14	6672.23	6397.04	0.955 ± 0.003	13
6788.83	5363.77	0.979 ± 0.008	10	5176.00	4734.77	0.963 ± 0.007	13	6536.51	5580.79	0.954 ± 0.025	11
6245.14	6185.79	0.978 ± 0.003	11	5821.15	5766.16	0.963 ± 0.012	11	7030.26	6071.33	0.954 ± 0.013	11
7559.43	5363.77	0.977 ± 0.004	12	7405.73	5705.12	0.963 ± 0.007	14	7203.63	7119.84	0.954 ± 0.006	10
7077.90	5900.59	0.977 ± 0.009	11	5262.44	5170.49	0.962 ± 0.022	10	7138.62	6713.79	0.954 ± 0.018	13
6269.89	6245.14	0.976 ± 0.003	12	6594.30	6268.59	0.962 ± 0.008	12	6250.87	5821.15	0.954 ± 0.016	10
7405.73	6268.59	0.976 ± 0.007	13	6414.08	5645.45	0.962 ± 0.021	10	7357.60	6834.58	0.954 ± 0.015	12
5176.00	4683.03	0.975 ± 0.004	14	7360.55	7138.62	0.961 ± 0.012	13	7136.79	6174.86	0.954 ± 0.014	10
6523.29	5769.91	0.974 ± 0.008	11	6770.17	6709.49	0.961 ± 0.005	14	6978.47	6792.51	0.953 ± 0.009	12
6622.84	6245.14	0.973 ± 0.005	10	6591.51	6323.94	0.960 ± 0.009	10	7077.90	6788.83	0.953 ± 0.008	14
6594.30	5775.91	0.972 ± 0.012	13	6993.12	6245.14	0.960 ± 0.005	11	6624.90	4429.33	0.953 ± 0.015	10
4969.12	4963.92	0.972 ± 0.006	14	5363.77	5236.27	0.960 ± 0.009	12	6523.29	6449.27	0.952 ± 0.008	13
7557.88	6607.10	0.971 ± 0.011	12	4984.78	4969.12	0.959 ± 0.010	12	6498.00	6071.33	0.952 ± 0.012	11
7557.88	5363.77	0.971 ± 0.010	10	6060.31	5363.77	0.958 ± 0.007	12	6139.95	5580.79	0.952 ± 0.025	12
6548.99	5404.58	0.971 ± 0.010	13	6795.26	5236.27	0.958 ± 0.006	13	5775.91	5363.77	0.952 ± 0.008	10
6803.35	6498.00	0.970 ± 0.008	12	6379.25	5821.15	0.958 ± 0.009	11	6108.06	5363.77	0.952 ± 0.006	11
7061.09	6792.51	0.969 ± 0.008	13	7322.07	5900.59	0.958 ± 0.013	11	6397.04	5953.32	0.952 ± 0.011	10
6245.14	6195.99	0.969 ± 0.003	12	7360.55	6834.58	0.958 ± 0.020	10	5900.59	5236.27	0.952 ± 0.013	12
6845.37	5795.21	0.969 ± 0.018	10	6643.43	4780.10	0.958 ± 0.010	10	6520.74	6245.14	0.951 ± 0.005	12
7581.47	5363.77	0.969 ± 0.006	12	7203.63	6978.47	0.957 ± 0.006	11	6841.61	6174.86	0.951 ± 0.008	10
7360.55	7357.60	0.968 ± 0.008	13	6845.37	6811.27	0.957 ± 0.023	10	6973.70	6245.14	0.951 ± 0.009	11
6425.68	6245.14	0.968 ± 0.005	11	5925.91	5494.10	0.957 ± 0.010	14	6594.30	5780.64	0.951 ± 0.010	13
6362.26	6245.14	0.968 ± 0.006	10	7832.88	6245.14	0.957 ± 0.005	12	7385.85	6978.47	0.951 ± 0.008	14
6594.30	6195.99	0.967 ± 0.007	13	6460.36	5982.77	0.957 ± 0.008	10	7832.88	6594.30	0.950 ± 0.010	13
6841.61	6337.99	0.967 ± 0.011	10	6841.61	6594.30	0.957 ± 0.011	11	6841.61	5609.82	0.950 ± 0.008	14
6944.62	6324.91	0.966 ± 0.005	13	6536.51	5508.33	0.956 ± 0.009	10

Table 5. The most frequently identified DIB bands among highly-correlated pairs featuring $15 > n > 9$.

DIB (Å)	Appearance/count
5363.77	15
6245.14	13
6594.30	8
6841.61	4
6498.00	4
5821.15	4

The bands 5780.64, 6203.58, and 6284.05 Å appear in multiple sets and are often well-correlated with the same DIBs. The high correlations between the DIBs in each of these sets accompanied by the high appearance rate of each DIB among the most correlated pairs (more than seven appearances in Tables 2 and 4) indicate that these three sets may constitute DIB families.

The consistently high correlations between the understudied 5363.77 Å and the more well-known DIBs (6613.74, 6203.58, 6195.99, 5780.64, 5705.12, and 5487.64 Å) is intriguing. The correlations between 5363.77 Å and both 5487.64 and 5780.64 Å were found to be slightly higher than the well-known correlation between 6613.74 and 6195.99 Å. Perhaps, these correlations imply that these lines share a common origin.

DIB pairs containing the bands 6613.74, 6203.58, 6195.99, 5780.64, 5705.12, and 5487.64 Å exhibit similar correlations obtained with different data sets. Friedman et al. (2011) derived a

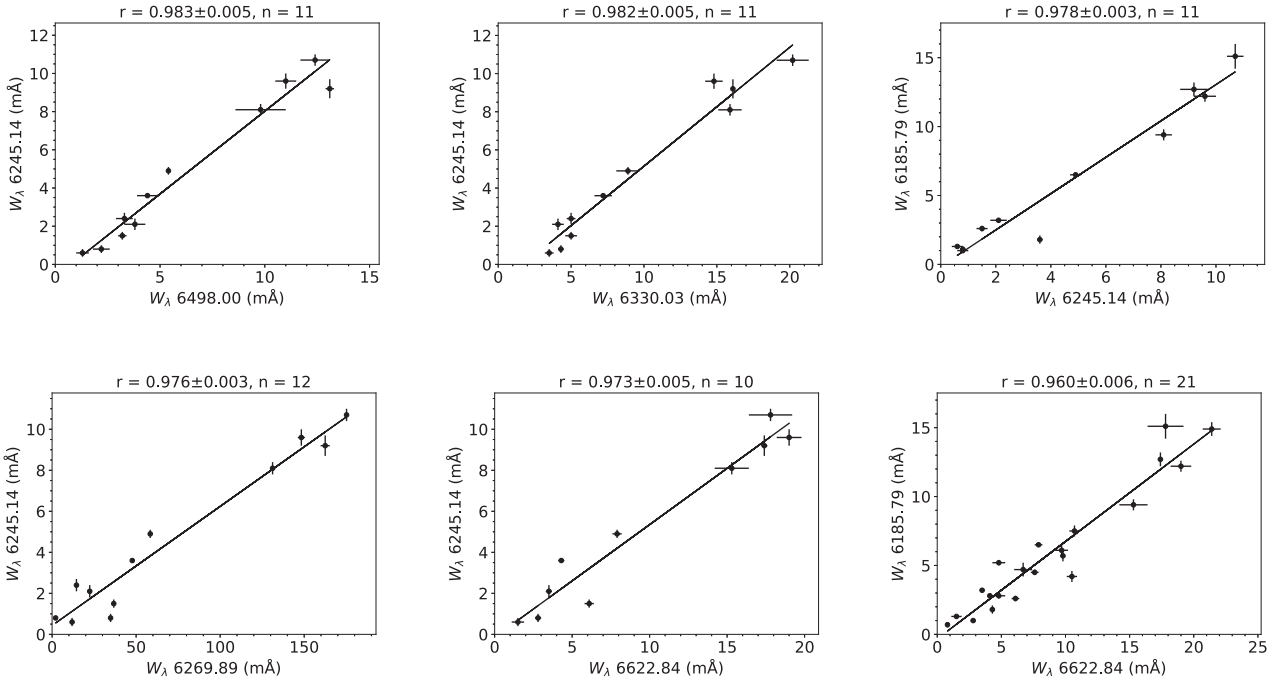


Figure 1. A subset of the highly correlated DIB pairs. In certain instances, the formal uncertainties are smaller than the representation of a given datum.

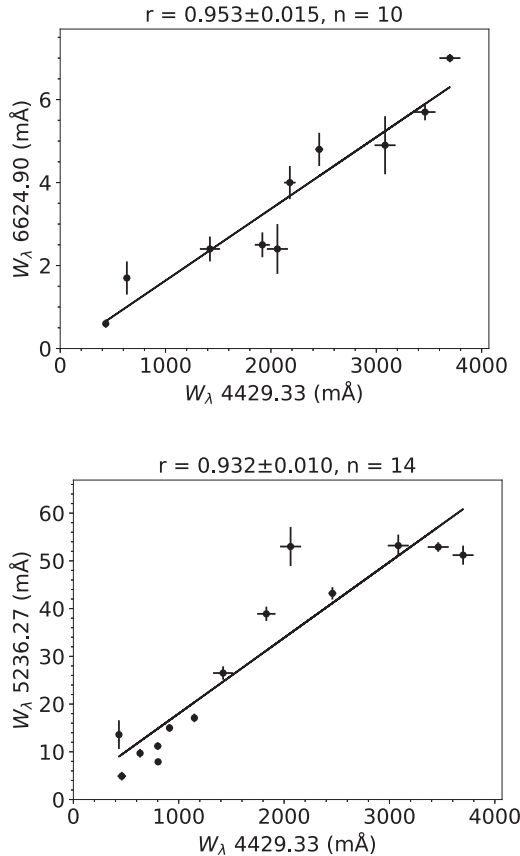


Figure 2. Correlation plots for 4429.33–6624.90 and 4429.33–5236.27 Å.

correlation coefficient of 0.99 for the pair 6196.0–6613.6 Å, whereas Bondar (2020) obtained 0.988. The value found here, for comparison, is 0.980 ± 0.001 . The pairs 5487.64–6284.05 and 5705.12–5797.18 Å have correlations that differ the most from previously reported values (Friedman et al. 2011). Marginal deviations from the coefficients may stem from sample size differences.

Most of the DIB pairs with lower correlations tend to involve at least one weak band (possibly both DIBs). The less studied DIBs (5363.77, 5609.82, 6245.14, and 6993.12 Å) are quite weak. Table 10 provides the average $EW/E(B - V)$ for certain DIBs (see also Table 2 in Fan et al. 2019). The highest uncertainty in the correlations among the three DIB sets, the correlation between 5363.77 and 5705.12 Å, was between the two weakest bands in that DIB set. In addition, among the highest uncertainties in Tables 2 and 4 was 0.025 between 6536.51 and 5580.79 Å, and the average $EW/E(B - V)$ of these DIBs are 8.04 and 3.13 mÅ mag^{-1} , respectively (Fan et al. 2019). Correlation analyses should consider weaker bands and avoid excluding potentially related DIBs. DIBs should not be excluded on the sole basis of high uncertainty values.

Uncertainties may be exacerbated by the inclusion of broad DIBs, since for example the continuum placement is challenging to identify (Fan et al. 2019). Consequently, caution is warranted when considering correlations involving such DIBs.

The bands 5780.64, 6203.58, and 6284.05 Å are each found in at least two of the three sets of DIBs identified here. The DIB 5780.64 Å is common to all three sets. Possibly, these three sets of DIBs have chemical similarities. Inspection of line profiles and investigation of possible carriers will be carried out, in the future.

High correlations are good indicators of DIBs with common origins. However, a correlation may appear to be low due to even a single outlier. Simple least-squares analysis does not account for such outliers, which may result in missing correlations between other DIBs. It is of interest to factor the sample size into the correlations; even if a correlation between two DIBs may not be

Table 6. Pearson r correlation coefficients and their uncertainties for a DIB set potentially tied to 4429.33 Å. The uncertainties were calculated through Monte Carlo simulations.

DIB (Å)	4429.33	4780.10	5236.27	5953.32	6009.60	6624.90
4429.33	1.000	0.940	0.932	0.938	0.915	0.953
	-	± 0.005	± 0.010	± 0.010	± 0.008	± 0.015
4780.10	0.940	1.000	0.900	0.901	0.877	-
	± 0.005	-	± 0.012	± 0.016	± 0.010	-
5236.27	0.932	0.900	1.000	0.934	0.878	-
	± 0.010	± 0.012	-	± 0.010	± 0.009	-
5953.32	0.938	0.901	0.934	1.000	0.904	-
	± 0.010	± 0.015	± 0.011	-	± 0.013	-
6009.60	0.915	0.877	0.878	0.904	1.000	-
	± 0.008	± 0.010	± 0.008	± 0.013	-	-
6624.90	0.953	-	-	-	-	1.000
	± 0.015	-	-	-	-	-

Table 7. The Pearson r correlation coefficients and their uncertainties for DIBs possibly linked to 5363.77 Å. The uncertainties were calculated through Monte Carlo simulations.

DIB (Å)	5363.77	5487.64	5705.12	5780.64	6203.58	6284.05
5363.77	1.000	0.981	0.956	0.980	0.964	0.950
	-	± 0.005	± 0.007	± 0.003	± 0.005	± 0.005
5487.64	0.981	1.000	0.959	0.976	0.983	0.977
	± 0.005	-	± 0.004	± 0.003	± 0.002	± 0.003
5705.12	0.956	0.959	1.000	0.980	0.972	0.972
	± 0.007	± 0.004	-	± 0.001	± 0.002	± 0.002
5780.64	0.980	0.976	0.980	1.000	0.986	0.978
	± 0.003	± 0.003	± 0.001	-	± 0.001	± 0.001
6203.58	0.964	0.983	0.972	0.986	1.000	0.990
	± 0.005	± 0.002	± 0.002	± 0.001	-	± 0.001
6284.05	0.950	0.977	0.972	0.978	0.990	1.000
	± 0.005	± 0.003	± 0.002	± 0.001	± 0.001	-

Table 8. The Pearson r correlation coefficients and their uncertainties for the set of DIBs associated with 5609.82 and 7224.16 Å. The uncertainties were calculated through Monte Carlo simulations.

DIB (Å)	5487.64	5609.82	5780.64	6203.58	6284.05	7224.16
5487.64	1.000	0.970	0.976	0.983	0.977	0.955
	-	± 0.004	± 0.003	± 0.002	± 0.003	± 0.003
5609.82	0.970	1.000	0.955	0.975	0.975	0.962
	± 0.004	-	± 0.004	± 0.003	± 0.003	± 0.003
5780.64	0.976	0.955	1.000	0.986	0.978	0.951
	± 0.003	± 0.004	-	± 0.001	± 0.001	± 0.001
6203.58	0.983	0.975	0.986	1.000	0.990	0.974
	± 0.002	± 0.003	± 0.001	-	± 0.001	± 0.001
6284.05	0.977	0.975	0.978	0.990	1.000	0.977
	± 0.003	± 0.003	± 0.001	± 0.001	-	± 0.001
7224.16	0.955	0.962	0.951	0.974	0.977	1.000
	± 0.003	± 0.003	± 0.001	± 0.001	± 0.001	-

especially high. The sensitivity of correlation coefficients to outliers may also explain slight differences between correlations found here and those found in previous works. An enhanced characterization of systematic uncertainties is needed to better assess outliers. That could be achieved in part by examining additional sightlines and using high-resolution spectra with high signal-to-noise ratio.

The current obstacle in analysing the suggestion that anionic hydrogen clusters may be DIB carriers is that multiple observed DIBs lie within the uncertainties of the current theoretical bands of the hydrogen clusters. For example, the cluster H_{13}^- has a calculated band at 5109.7 Å, but three bands, 5100.92, 5110.86, and 5117.63 Å

lie relatively close to this. However, with higher levels of quantum chemical calculations and smaller uncertainties, it may be possible to construct a better understanding of these clusters' relation to DIBs and make more robust conclusions (supportive or otherwise). The correlations between lines proposed to be associated with anionic hydrogen clusters suggested in Huang et al. (2019) was also investigated. Sufficiently strong correlations could not be found between the associated DIBs to draw a definitive conclusion at this time. Once more refined calculations are available for the anionic hydrogen clusters, their viability as a DIB carrier will be further investigated.

Table 9. The Pearson r correlation coefficients and their uncertainties for the set of DIBs centred around 6269.89 and 6993.12 Å. The uncertainties were calculated through Monte Carlo simulations.

DIB (Å)	5780.64	6195.99	6269.89	6613.74	6993.12
5780.64	1.000	0.974	0.961	0.948	0.959
	-	±.001	±.001	±.001	±.001
6195.99	0.974	1.000	0.975	0.980	0.972
	±.001	-	±.001	±.001	±.001
6269.89	0.961	0.975	1.000	0.972	0.984
	±.001	±.001	-	±.001	±.001
6613.74	0.948	0.980	0.972	1.000	0.952
	±.001	±.001	±.001	-	±.001
6993.12	0.959	0.972	0.984	0.952	1.000
	±.001	±.001	±.001	±.001	-

Table 10. Mean $EW/E(B - V)$ for a subsample of the DIBs examined (an abridged version of table 2 from Fan et al. 2019).

DIB (Å)	$\langle EW/E(B - V) \rangle$ (mÅ mag ⁻¹)
5609.82	12.34
5363.77	17.66
6195.99	47.06
6993.12	64.72
6269.89	70.16
5705.12	90.65
5487.64	108.48
6203.58	157.31
7224.16	162.17
6613.74	185.14
5780.64	398.84
6284.05	958.09

4 CONCLUSIONS

The Fan et al. (2019) database of 557 DIBs was analysed. Here, 231 pairs of highly correlated ($r > 0.95$) DIBs were identified, and are highlighted in Tables 2 and 4. 12 DIBs, 5363.77, 5487.64, 5609.82, 5705.12, 5780.64, 6195.99, 6203.58, 6269.89, 6284.05, 6613.74, 6993.12, and 7224.16 Å, are found to have high correlations and belong to strongly inter-correlated DIB sets. Three such sets are newly identified, with five of the 12 DIBs being understudied. One set consists of 5363.77, 5487.64, 5705.12, 5780.64, 6203.58, and 6284.05 Å, another consists of 5487.64, 5609.82, 5780.64, 6203.58, 6284.05, and 7224.16 Å, and the other contains 5780.64, 6195.99, 6269.89, 6613.74, and 6993.12 Å. A set of correlated lines including the famed 4429.33 Å band is also identified for the first time.

In addition, previously identified correlations between certain DIBs were bolstered. The bands 5780.64, 6203.58, and 6284.05 Å are frequently found among the most correlated DIBs. The same DIBs tend to be part of the most highly correlated pairs, particularly 5363.77, 5780.64, 6195.99, 6203.58, and 6245.14 Å with more than

10 appearances among the highest correlation pairs. Correlations between the band 4429.33 Å and the other bands in this set, particularly 4780.10, 5236.27, 5953.32, 6009.60, and 6624.90 Å, are promising and warrant further analysis.

The correlations discussed provide an indication of possible relations between DIBs, as supported by numerical evidence. A future investigation concerns inspecting the spectral profiles for similarities. Furthermore, analysing a separate or larger spectral database is desirable.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. SeyedAbdolreza Sadjadi (University of Hong Kong) for helpful comments. The authors are also grateful to the anonymous Reviewer who provided helpful feedback. The Python libraries pandas (McKinney 2010; Reback et al. 2020), Matplotlib (Hunter 2007), and Scipy (Virtanen et al. 2020) were utilized in this study and are also acknowledged for their usefulness and open availability. The Natural Sciences and Engineering Research Council of Canada (NSERC), Canada Foundation for Innovation (CFI), and Mount Saint Vincent University are acknowledged for funding.

DATA AVAILABILITY

The data underlying this article are available in the Centre de Données astronomiques de Strasbourg (CDS), at <https://cdsarc.unistra.fr/viz-bin/cat/J/ApJ/878/151> [10.26093/cds/vizier.18780151]. The data found in Tables 2 and 4 are available in the article. The data found in Tables 2 and 4 will also be available in CDS.

REFERENCES

- Bondar A., 2020, *MNRAS*, 496, 2231
- Cami J. et al., 1997, *A&A*, 326, 822
- Fan H. et al., 2019, *ApJ*, 878, 151
- Friedman S. D. et al., 2011, *ApJ*, 727, 33
- Galazutdinov G. A., Krelowski J., 2017, *AcA*, 67, 159
- Galazutdinov G. et al., 2020, *AJ*, 159, 113
- Huang L. et al., 2019, preprint ([arXiv:1912.11605](https://arxiv.org/abs/1912.11605))
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Krelowski J., Walker G. A. H., 1987, *ApJ*, 312, 860
- Linnartz H. et al., 2020, *J. Mol. Spect.*, 367, 111243
- McCall B. J. et al., 2010, *ApJ*, 708, 1628
- McKinney W., 2010, in van der Walt S., Millman J., eds, *Proceedings of the 9th Python in Science Conference, Data Structures for Statistical Computing in Python*. Austin, Texas, p. 56
- Moutou C. et al., 1999, *A&A*, 351, 680
- Reback J. et al., 2020, Zenodo, pandas-dev/pandas: Pandas 1.0.2. Available at: <https://zenodo.org/record/3708035>
- Virtanen P. et al., 2020, *Nature Methods*, 17, 261
- Xiang F., Liu Z., Yang X., 2012, *PASJ*, 64, 31

APPENDIX A: ADDITIONAL FIGURES

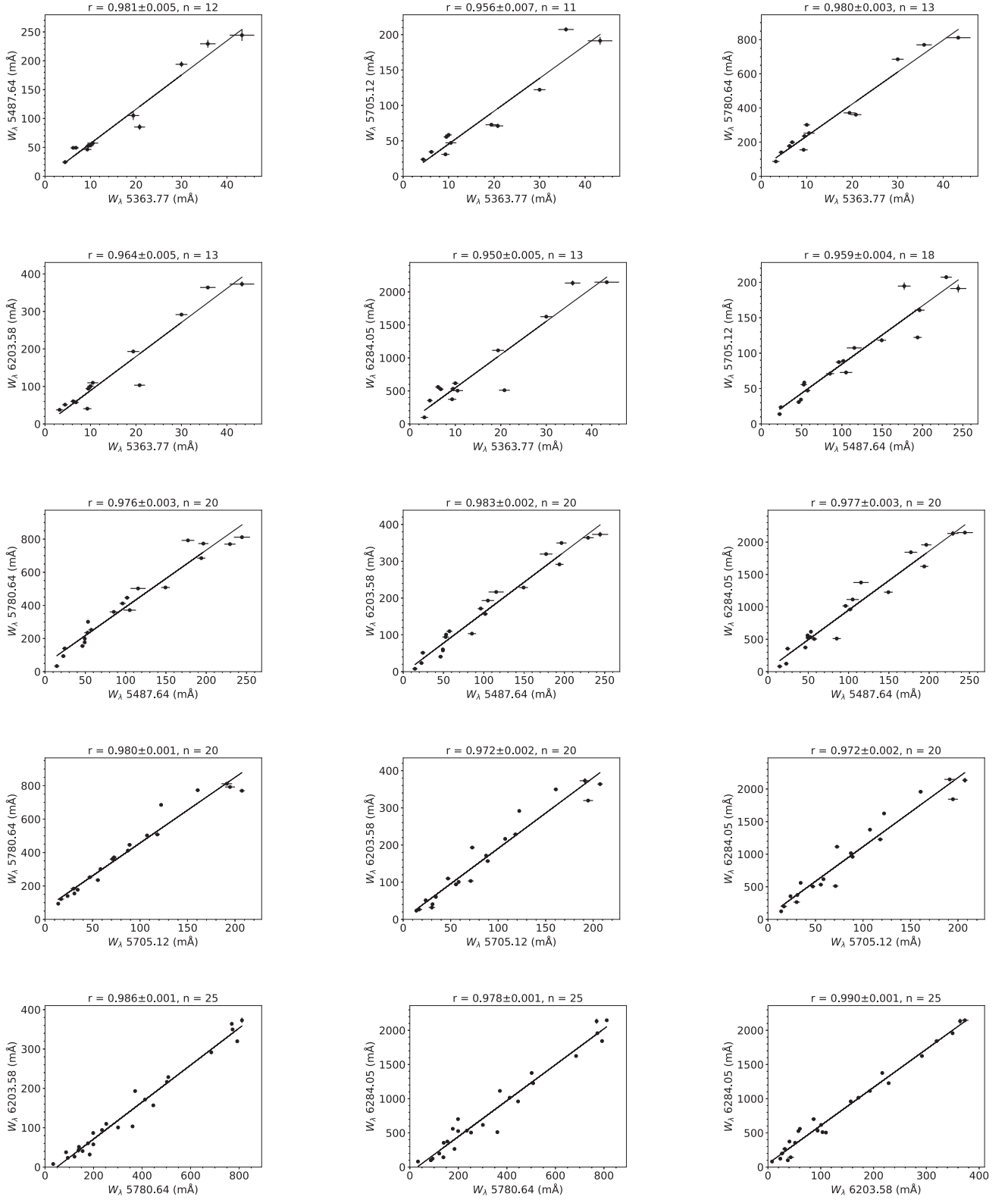


Figure A1. Correlation plots for pairs in the first identified DIB set.

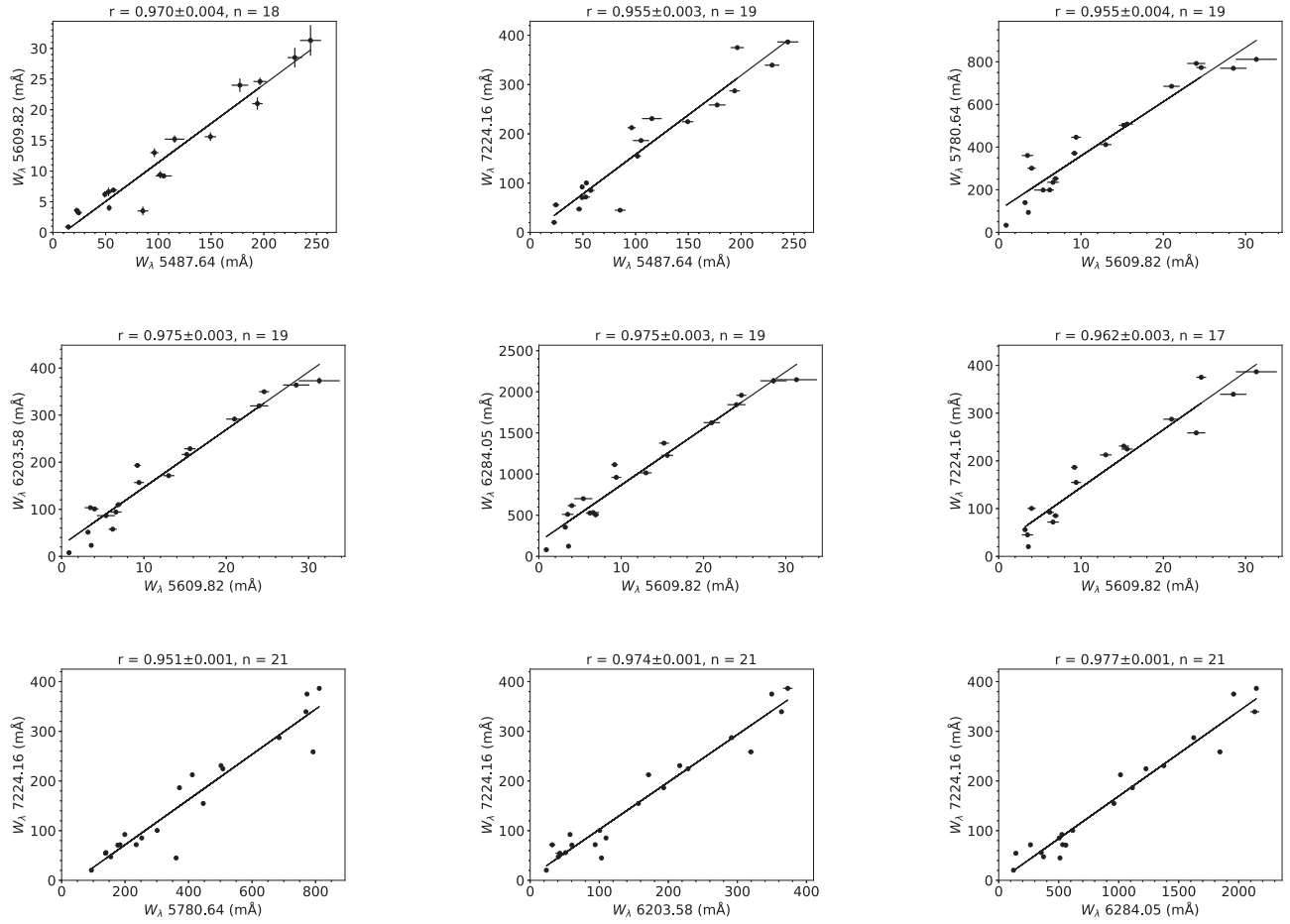


Figure A2. Correlation plots for pairs in the second identified DIB set.

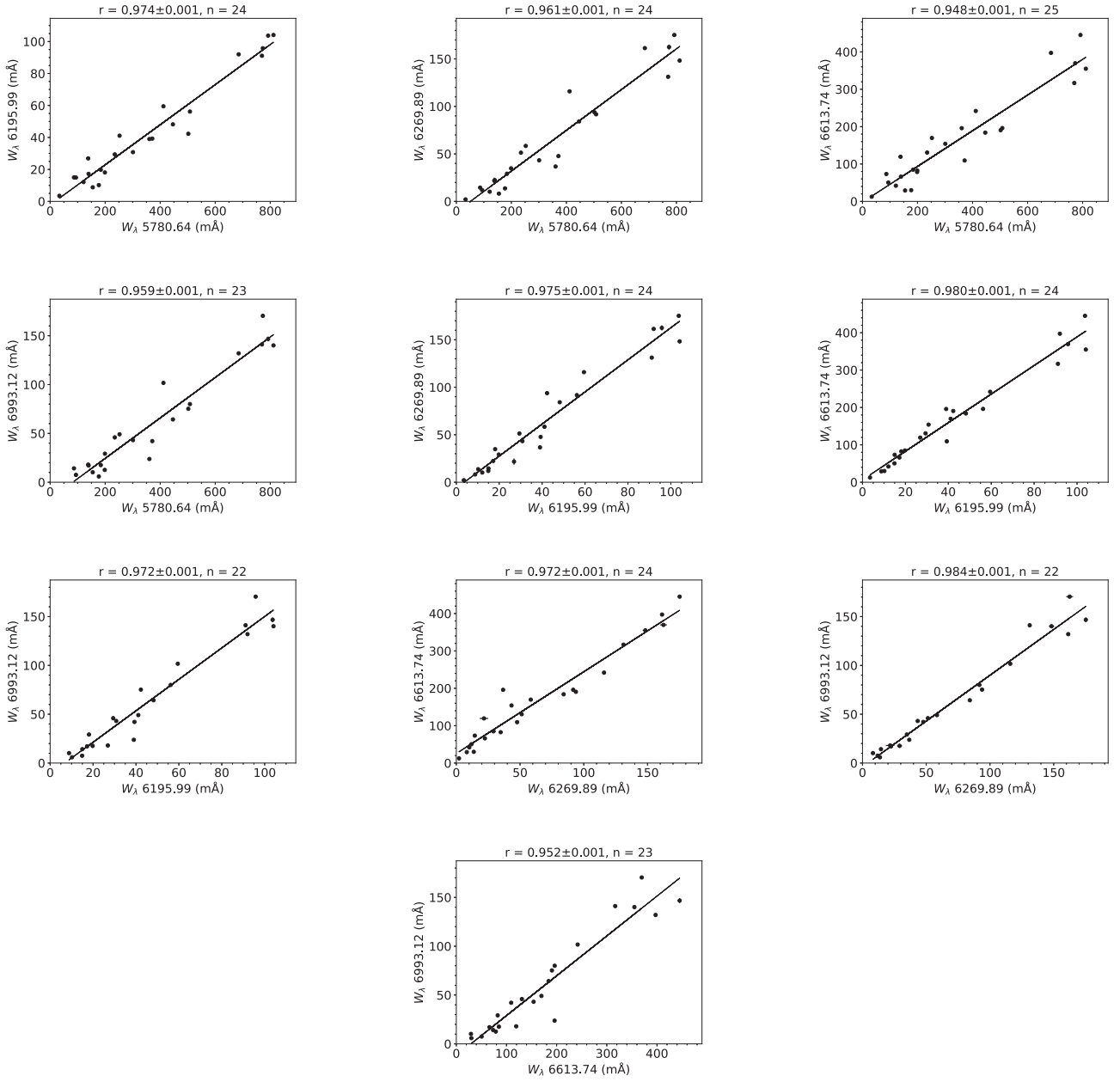


Figure A3. Correlation plots for pairs in the third identified DIB set.

This paper has been typeset from a \LaTeX file prepared by the author.