

Species diversity and spatial distribution of CL/VL vectors: Evidence for variation in gene expression of immunogenic SP15 and LelF proteins due to bioclimatic characteristics and physiological status of sand flies, Iran

Ali Bordbar

Institute Pasteur of Iran

Parviz Parvizi (✉ parpparvizi@yahoo.com)

Pasteur Institute of Iran

Research

Keywords: CL and VL vectors, immunogenic proteins, spatial distribution, GIS modeling, biodiversity, SP15 and LelF genes expression, bioclimatic regionalization

Posted Date: February 26th, 2020

DOI: <https://doi.org/10.21203/rs.2.24563/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background Main approaches of this investigation were climate regionalization to recognize the spatial distribution of cutaneous/visceral leishmaniasis (CL/VL) vectors as risk-maps using ArcGIS modeling system, evaluation of species biodiversity, assessing bio-climate (BC) effect on expression plasticity of genes possessing vaccine properties isolated from wild-collected sand flies. **Methods** The Inverse Distance Weighting (IDW) interpolation method was used to obtain accurate geography map using ArcGIS10.3.1 in closely-related distances. Species diversity was calculated based on Shannon diversity index using Rv.3.5.3. Expression fold change of SP15 and LelF genes was evaluated using cDNA synthesis and RT-qPCR analysis. **Results** Three BC zone were identified in northeast of Iran. *Phlebotomus papatasi* were abundantly captured in all BC regions and the frequency was more in plains areas of mountainous BC as hot spots of CL. VL vectors were more prevalent in spatial cluster of Mediterranean BC. Semi-arid BC was identified as a major contributing factor to up-regulate SP15 salivary gene expression ($P = 0.0050$, $P < 0.05$), and Mediterranean BC had considerable effect on up-regulation of LelF- Leishmania gene in gravid and semi-gravid *P. papatasi* population ($P = 0.0109$, $P < 0.05$). **Conclusions** The diversity and spatial distribution of CL/VL vectors associated with BC regionalization obtained in our research provide epidemiological risk maps and establish more effectively control measures against leishmaniasis. Oscillations in gene expression indicate that each gene has its own features, which are profoundly affected by bioclimatic characteristics and physiological status of sand flies. It is essential to consider BC factors affecting regulatory regions of environmentally responsive loci for genes used in vaccine designing.

Background

Leishmaniasis is an epidemic-prone infectious and neglected tropical disease (NTD) caused by obligate intra-macrophage protozoa of the genus *Leishmania*. The *Leishmania* parasites are transmitted by the bite of infected female sand flies to humans, which resulting in an extensive range of clinical manifestations mainly from self-healing cutaneous leishmaniasis (CL, the most common) to progressive fatal visceral leishmaniasis (VL) [1, 2]. Geographical spread of the disease is extended throughout widespread territories and countries (98 countries) in all over the world, except the Australia, the Pacific Islands, and Antarctica [3]. More than 95% of new CL cases occur in six countries including Afghanistan, Algeria, Brazil, Colombia, Iran, Iraq and the Syrian Arab Republic [2]. Iran is one of the most important foci of leishmaniasis in all around the world, which is one of the major public health problems in more than half of the provinces [4]. Iran's geographic location provides an appropriate climate and landscape habitat for different species of rodents and dogs as main reservoirs and sand flies as principle vectors [5].

In northeast of Iran, CL and VL were reported from all counties and five out of eight counties of Northern Khorasan Province, respectively [6]. It is essential to investigate the ecology, biodiversity, adaptation to habitats, and distribution of sand flies, according to environmental variables involved in transmission, therefore control measures and spatial modeling of infectious diseases can be obtained using the Geographic Information System (GIS)-based survey [5, 7, 8]. Biodiversity includes intraspecific (genetic polymorphism), interspecific, and ecosystem diversity. Evaluation of biodiversity measurements can be achieved based on spatial scale [9] as followings: alpha diversity (α) (the number of species within a community like a natural habitat), beta diversity (β) (interspecific differences with which we can rapidly compare the change rate of diversity at different habitats), and gamma (γ) or regional diversity (diversity among ecosystems). Calculating the diversity indices and comparing their results is a useful method for recognizing the ecological status of sand flies. Hence, determining biodiversity indices of sand flies in different topographic locations and the study of correlation between climatic elements and incidence of leishmaniasis can help us developing an effective prevention approach against vector-borne diseases [10, 11, 12]. Despite the worldwide distribution of leishmaniasis and also its significant disease burden, currently there are no commercially available human leishmaniasis vaccines delivering the necessary level of protection. Due to the lack of efficacious vaccines to treat leishmaniasis, studying the vectors is an important component of the effort to control leishmaniasis [13, 14]. Prior to the use of antigenic proteins specific to the genus *Leishmania* or exploited from sand fly saliva as a vaccine target, the genetic and expression changes of salivary and *Leishmania* proteins firstly needs to be thoroughly assessed in natural-field habitats.

In our recent ongoing research, bioinformatics analyses demonstrated that the vaccine comprising the combination of truncated forms of salivary protein from *P. papatasi* (SP15) and *Leishmania* eukaryotic initiation factor (LelF) namely SaLeish, has an appropriate traits to increase both innate and specific cellular immune responses with acceptable population coverage in different endemic areas of the world [15]. The protein PpSP15 was the first identified salivary protein, and could be used in a sand fly salivary protein-based vaccine strategy as saliva-driven protective immunity against vector-transmitted leishmaniasis [16, 17]. Indeed, the immunization obtained by SP15 from *P. papatasi* and its orthologs in *P. dubosqi* (PdSP15) was sufficient for protection through specific delayed-type hypersensitivity (DTH) response with a Th1 profile in Non Human Primates (NHP) and humans [18, 19]. It has been indicated that SP15 of *P. papatasi* salivary has the ability of stimulating the dendritic cells (DC), the potent antigen presenting cells, and therefore initiating the saliva-mediated immunity conferring the induction of rapid sand flies saliva-specific Th1 immune response [16, 20]. Another protein selected to be examined was LelF, which was obtained from natural collected-sand fly of Northern Khorasan province. LelF is homologous and highly conserved among *Leishmania* spp., which can be used as a natural adjuvant for antigen-based vaccine strategy and results in inducing the early production of LelF-specific Th1 cytokines, IL-12 and IFN γ in PBMC of both leishmaniasis patients and normal individuals [21, 22]. The LelF protein has previously been used as leishmaniasis prophylactic vaccine (Leish-111f) that elicited an increase in CD4⁺ T cells and thus indicating a predominant Th1-type immune response [23]. There are few studies considering the eco-epidemiological effects according to bio-climate (BC) regionalization on the expression of sand fly salivary proteins as well as their *Leishmania* within at peak activity of sand flies in leishmaniasis endemic regions. Besides, a number of studies have merely investigated the effect of

environmental response on the evolution of gene expression, regardless of the precise determination of the locations affecting the expression of genes of target associated with GIS modeling system [24].

The aims of this investigation were 1) determining the spatial distribution of potential and principle vectors of CL and VL based on BC regions in ArcGIS tool as hazardous maps in northeast of Iran, 2) to regionalize the ecological habitats of sand flies as bio-climate zones according to the environmental factors using integrative approaches, species diversity assessment and richness, 3) similarity analyzing in sand fly populations collected in different ecotopes based on Shannon-Wiener index and Jaccard coefficient using R, 4) evaluating different gene expression of immunogenic proteins (SP15 and LelF) in wild-caught *P. papatasi* salivary gland and from natural *Leishmania*-sand fly in relation to the effects of bio-climate factors and physiological statuses of sand flies appropriate for protein-based vaccine design.

Methods

Specifications of study area

Integrative methods were used to evaluate eco-epidemiological parameters such as bioclimatic change, landscape fragmentation, and ecological niche considering vector domiciliation and their distribution in different locations of Northern Khorasan foci. The Northern Khorasan province is located in northeast of Iran and from the epipaleolithic period onwards was substantially inhabited by nomadic hunter-gathers. Geographic location of studied area (GPS coordinates) is between latitudes of 36°34' -38°17' N and longitudes of 55°52' -58°20' E, and has an area of about 28,434 km² including eight counties: Bojnord (the capital city of province), Shirvan, Esfarayen, Maneh-o Samalghan, Jajarm, Faruj, Garmeh, Raz-o Jargalan (Fig. 1). This province is bordered Turkmenistan from north, with Khorasan-Razavi province from south and east, with Golestan province from west, and with Semnan province from the southwest (Fig. 1). Northern Khorasan district is generally regionalized into two fractions as: Mountainous areas, and flat regions with low elevation. The Kope-Dagh to the north and Aladagh, extending of the Alborz Mountains to the south, are two main ranges with fertile plains between the mountains, which determine the favorable agricultural conditions and pastoralism. Regarding various weather conditions, this area has a thriving culture and vegetation. Regions of Bojnord, Shirvan, Faruj, and Esfarayen are cold mountainous whereas western districts of Maneh-o Samalghan and Raz-o Jargalan counties are temperate and forested. Parts of lush green forest of Golestan National Park (one of the oldest national parks in the world) are located in Northern Khorasan Province. However, this province has also desert areas such as Jajarm, where there is a wide range of temperature oscillations (cold winters, and relatively hot and dry weather in summer) (Additional file 1: Figure S1a, b, c). Furthermore, significant rainfall amount occurs at different intensities in the province. Due to vast mountainous areas, the Northern Khorasan zone is mainly known to have a cold to temperate mountainous region. The provincial climate is affected by longitude and latitude, and altitude. Also, three basic types of air masses bring variety of weather conditions to the Northern Khorasan area, including (1) Western humid air masses that enters the province in early autumn and lasts until late spring, and the peak of its activity during winter season is the main cause of rainfall in the province (Additional file 1: Figure S1b), (2) Siberian cold air masses, which affects the province from early autumn to early spring causing a significant temperature reduction (Additional file 1: Figure S1c), and (3) a warm dry air mass, which enters to the province from the south during summer, and increased the aridity and temperature (Additional file 1: Figure S1d).

Model variables and climatic regionalization

The De Martonne aridity index which is used to determine the regional climate based on annual temperature and rainfall is calculated as follows: $I = P/T + 10$ where (I) stands for De Martonne aridity index, P stands for annual rainfall in millimeters, and T stands for the annual temperature in °C (Fig. 1, Additional file 1: Figure S1). Also, climatic regionalization was carried out as earlier described [5] by the use of Principal Component Analysis (PCA) as multivariate statistical method to minimize sources within the variation of the group. Then, Clustering Integration Method (CIM) was applied to precisely improve the classification accuracy. The southern cities have higher temperatures compared to northern regions, due to their different altitudes from sea level, entrance of various air masses, and different longitude and latitudes (Additional file 1: Figure S1). The rainfall levels were also varied in different regions of the province. The maximum and minimum rainfall is observed in spring (March) and summer (June), respectively. Due to highly variability of sand fly species, even in closely distance from its sampled location, we utilized the Inverse Distance Weighting (IDW) deterministic interpolation using ArcGIS 10.3.1 to specify the collected sample points in terms of sand fly dispersion and targeted protein expression rates (SP15 and LelF) in local scale levels. IDW is a flexible spatial interpolation method, which estimates unknown values along with specifying the search distance, closest points, power setting and, barriers [25]. To obtain the geography map in GIS, ordinary IDW parameters related to the accuracy of the values at each point was calculated with respect to distance: (see Formula 1 in the Supplementary Files)

where Z_0 is the estimation value of variable z in point l , z_i is the sample value in point l , d_i is the distance of sample point to estimated point, N is the coefficient that determines weigh based on a distance, n is the total number of predictions for each validation case. The Jaccard's similarity coefficient (JSC) is the simplest data, which compares the similarity for two sets of data between different locations of communities. This index always gives a value with a range from the lowest (0%) to the highest similarity (100%) between two populations [26].

Site selection

This cross-sectional study was performed at the peak of sand fly activity from mid-June to the late of September (2017-2018) in the most susceptible biotopes to leishmaniasis where sand flies' habitats include rural (vegetation, plain and desert fields, and animal shelters), sylvatic-

habitats of reservoir hosts (including rodents and dogs) and urban regions (human dwellings). 17 main sites along with surrounding villages were selected from different locations in Northern Khorasan province (Fig. 1, Table 1).

Collecting the Sand flies

Considering geographical locations (altitude, latitude, and longitude) and weather conditions of the counties (Fig. 1, Additional file 1: Figure S1a, b, c, d) as well as during the peak seasons of sand flies activity (according to the incidence of leishmaniasis in each of these regions), trapping was carried out using 120 sticky oil papers, four CDC light traps and, 70 funnel traps. The catching time was from five to 10 days before sunset until sunrise at each location. Dead sandflies were removed from castor oil papers using fine brush and, were put into sterile tubes containing 96% alcohol, and then in order to identify the species, they were kept at -20 °C for future morphological and/or molecular experiments. Blood-fed, semi-gravid, and gravid female sand flies were collected alive by the use of aspirator, funnel traps and CDC light traps, and after that were immediately moved into suspended net cages using mouth aspirator. To provide paramount protection of captured sand flies, a piece of 30% sugar-soaked cotton pad was placed on the top panel of holding cages, and were then covered with plastic bags containing a moist sponge to avoid from losing humidity and/or excessive heat. Field-caught sand flies were first euthanized in water, dissected, and were then identified to species based on morphological characteristics of the head and terminal genitalia under compound microscope (400×) using a systematic key [27, 28].

RNA extraction from salivary gland and *Leishmania* in wild-caught sand flies

Expression pattern of SP15 and LelF proteins isolated from salivary glands and *Leishmania* protozoans of *P. papatasi* were incriminated in different locations of studied areas during the peak of sand fly activity from June to mid-September. Physiological status of sand flies was considered to determine SP15 gene expression level. Therefore, due to fresh-fed sand flies indicated a significant up-regulation of SP15 compared to un-fed ones [29], we selected only blood-fed sand flies to investigate the pattern of expression for SP15 transcript in different eco-epidemiologic conditions of Northern Khorasan province. After accomplishing definitive identification of female *P. papatasi*, salivary glands in addition to their head accessories were dissected using sterile forceps, and then were immediately re-suspended in 50 µl of RNA-zole® RT solution (MRC, OH, USA), homogenized, and stored at -20°C. To assess the proper expression of SP15, pulled off parts of 10 sand flies were pooled for total RNA isolation (Jena Bioscience, Germany).

To evaluate the expression of LelF, sand fly-derived metacyclic promastigotes were distinguished by morphology under loop microscope. As parous sand flies (flies that lay eggs) play crucial role in transmitting disease; fresh-fed, gravid, and semi-gravid sand flies were subjected to microscopic examination. Then, the head along with the thoracic midgut forefront which is located in the vicinity of stomodeal valve (SV), and SV area of each infected sand fly were dissected based on the physiological status. Also, parasite isolation from proboscis extending to stomodeal valve of sand fly was not exempt from technical difficulties. Promastigotes obtained from aforementioned area of *P. papatasi* were counted using Neubauer chamber. The numbers of parasites isolated from each sand fly was approximately 10⁴ promastigotes. RNA Purification was implemented for high yield isolation of total RNA from *Leishmania* parasites of 10 sand flies with specified status utilizing silica-gel membrane adsorption kit (Jena Bioscience, Germany, <http://www.jenabioscience.com>) and stored at -80°C until required. RNA quality and quantity were determined using Nanodrop (BioTek® Synergy™ HTX Multi-Mode, USA).

cDNA synthesis

Complementary DNAs (cDNA) were synthesized as a template for performing quantitative Real-Time PCR (qRT PCR) using the Easy™ cDNA Synthesis Kit (Parstous Biotech, <http://www.parstous.com/easy-cdna-synthesis-kit.html>, Mashhad, Iran), with respect to the manufacturer's instructions. Briefly, 8 µl of template RNA from each sample was mixed with kit components in RNase-free tube. After quick vortex, prepared mixture was incubated for 10 minutes (min) at 25°C, and then was incubated for one hour at 47°C following stopping the reaction by heating for 5 min at 85°C, and chilling on the ice for 10 min. To perform PCR or quantitative PCR, the finished RT reaction can be added up to 1/5 of the final PCR volume.

Quantitative Real-time polymerase chain reactions (RT-qPCR)

Real-time PCR reactions were performed to quantify the total amount of SP15 and LelF transcript expression from salivary gland and *Leishmania* parasites of *P. papatasi* sand flies using Maxima SYBRgreen and Rotor-Gene Q instrument (Qiagen, Germany). The qRT-PCR reactions were set up in duplicates with a total volume of 10 µl, containing 5 µl SYBR green master mix (Fermentas, UK), 0.5 µl each forward and reverse primer (0.8 µM final concentration), 1.5 µl template of cDNA and 2.5 µl Ultra-Pure DNase/RNase-Free water (Invitrogen) in 0.1 ml capillary tubes. Initial reaction was incubated for 10 min at 95 °C, followed by 40 cycles of 95 °C for 15 s (sec), annealing for 30s at 60 °C and extension for 30s at 72 °C. Fluorescent emission was measured at 530 wavelength for SYBR Green dye at the end of the elongation cycle. Expression of SP15 of salivary gland and LelF of *Leishmania* gene fragments were assessed in each one of 10 sand fly groups. Each reaction was repeated four times for each gene (in duplicate in two different runs). Quantitative Real-Time PCRs were performed with a 159 bp fragment of the SP15 cDNA using specific primer pairs for SP15 (forward: SP58F, 5'-TGCATTCGCCATTGCAAATATCAG-3', reverse: SP194R, 5'-AGCACATTCGTGCATAATTTTC-3'), and LelF (forward:

LF1014F, 5'-CGTCATCAACTTCGACCTGCC-3', reverse:

LF1153R, 5'-TCGATCTGCGTGTGGTAGTG-3') with a 159 bp fragment for the LelF cDNA. α -tubulin gene was used as a housekeeping load control mRNA to the qPCR reactions primers Tub-P24F and Tub-P24R in *P. papatasi* sand fly [30]. All PCR experiments were followed by melting curves to check the amplicon specificity.

Different expression ratios for salivary protein and *Leishmania* genes (SP15 and LelF) were calculated and displayed as fold changes over a control, using the $2^{-\Delta\Delta CT}$ method [31]. The fold changes were calculated by the expression $2^{-\Delta\Delta CT}$, where $\Delta\Delta CT = \Delta CT$ (sample) - ΔCT (calibrator), $\Delta CT = \Delta CT$ (sample) - ΔCT (alpha tubulin gene), C_T = cycle at the statistically significant increase in the emission intensity over the background. The calibrator was represented by the average expression (mean ΔCT) of the newly emergent un-fed nulliparous laboratory-reared female sand flies [30]. Moreover, negative controls were considered without the templates. Fold changes were calculated for each sample in comparison with the calibrator.

Data analysis

Statistical analyses were carried out using the software GraphPad Prism v. 5.01 (GraphPad Software, Inc). Non-parametric KruskalWallis tests were performed to compare the multiple datasets with Mann-Whitney *U* tests, for accomplishing 2-way pairwise comparisons between data sets when the results for Kruskal-Wallis test were statistically significant, or for comparisons when only two data sets were present. The D'Agostino-Pearson normality test was computed the **skewness and then the kurtosis**, to quantify how far the distribution is from Gaussian in terms of asymmetry and shape. After that, it calculated how far each of these values differs from the value expected with a Gaussian distribution, and computed a single *P* value from the sum of these discrepancies. Species biodiversity used in this investigation is affected by two factors as: (1) number of species and (2) their abundance of distribution [26]. Species diversity was calculated based on Shannon-Wiener index, species richness and also evenness index using the Ade4 package of R v.3.5.3 software to estimate the biodiversity of sand fly species in the study areas. Species diversity is a measure of diversity within an ecological community, which encompasses both species richness and evenness of species abundance. Species richness is the number of species in a specific area or environment, whereas, species evenness refers to how close each species are in number in the similar location. Evenness is species abundance distribution and the more even this distribution is the greater the species evenness will be [32]. Computation of these indices is shown in Table 3. Also, estimation by similarity was calculated between different communities using Jaccard's similarity coefficient within different localities collection (Table 4).

Results

Sand fly distribution and species collections

In this study, 12 species of sand flies belonging to 8 species for genus *Phlebotomus* and 4 species for genus *Sergentomyia*, were identified (Table 1, Additional file 2: Table S1). A total of 1167 sandflies (283 males and 884 females) were collected utilizing CDC light traps, sticky traps, aspirators and, funnel traps from three climate zones including six districts (Table 1, Additional file 2: Table S1, Additional file 3: Figure S2).

We analyzed the expression of SP15, protein of salivary glands, from blood-fed of uninfected female in 10 individual *P. papatasi* sand flies from each location at each period of the season (early, mid, or late) (Figs. 6, 7, 8, Table 2, Additional file 4: Table S2), and also the LelF expression was quantified from *Leishmania major* isolated from infected female of *P. papatasi* sand fly with fresh blood-fed, gravid, and semi gravid status (Figs. 9, 10, 11, Table 2, Additional file 5 Table S3). Four principal climatic elements of studied region (wind, annual precipitation, temperature, and humidity) were considered to determine bioclimatic classification and distribution of sand fly species of Northern-Khorasan province (Figs. 1, Additional file 1: Figure S1). Among aforementioned factors, annual temperature, humidity and rainfall have the greatest effects on the sand flies lives. After evaluating the impact of principle climatic factors on climatic variables, three regions were classified as homogeneous climate zones (Table 1, Additional file 2: Table S1 (stations sheet), Additional file 3: Figure S2). The regions are as follows: (1) The Mediterranean with spring rains area in northwestern including Razo- Jargalan, and Maneh-o Samalqan (Table 1, Additional file 1: Figure S1, Additional file 2: Table S1). (2) The cold-mountainous area in southwestern including northern parts of Garmeh, in central including Bojnord, and in eastern including Shirvan, Faruj, and Esfarayen. (3) The low rain fall and warm-semi dry area including Jajarm, southern and western parts of Garmeh and Esfarayen, respectively (Additional file 1: Figure S1).

The data of CL and VL vector presence (*P. papatasi*, VL species: *chinensis* group, *major* group, *P. kandelaki*, *Paraphlebotomus* species: *P. sergenti*, *P. caucasicus*, *P. ansari*, *P. alexandri*, *Sergentomyia* species: *S. sintoni*, *S. dentata*, *S. sumbarica*, *S. baghdadis*) were imported to ArcMap, which transfers the raw data into visualized information to be used as a hazard map for ZCL, ACL, and VL transmission in each bioclimatic zones of Northern Khorasan province (Figs. 2, 3, 4, 5, Additional file 2: Table S1). In the genus *Phlebotomus*, *P. papatasi* were in all climate regions ($n=800$ specimens: 68.55%). Regarding, *P. papatasi* was the most abundant one in spatial cluster of mountainous regions particularly in Baba Aman, Quch Qaleh, and Paygho where human dwellers are almost in all ecotopes, ranging from 11.25% to 18% (Fig. 2, Table 1), and the least abundant one of them was observed in Gholaman and Darkesh of Mediterranean region, and the farthest eastern part of mountainous regions (Fig. 2). *P. alexandri*, proven vector of VL and suspected vector of CL, was the second-most widespread species caught in mountainous climate zone (Fig. 3, Table 1, Additional file 2: Table S1 (xy sheet)). From 153 *P. alexandri*, 60.79% ($n=93$ specimens) were caught from rodent burrows' ecotope in sylvatic habitats, while 39.21% *P. alexandri* ($n=60$ specimens) were found in peridomestic animal shelters. *Paraphlebotomus* species were found only in spatial cluster of mountainous and Mediterranean bio-climate (BC) zones (Fig. 3, Table 1). *Sergentomyia* genus was prevalent almost in all three bio-

climate areas (Fig. 4, Table 1). However, *Sergentomyia* was more prevalent only in spatial cluster of Mediterranean and mountainous BC zones especially more for Gholaman, Qareh Bashloo and, Quch Qalleh-Olya with the mean of range between 10.26-16.33% (Fig. 4, Table 1). *S. baghdadis* and *S. sintoni* were predominant species of *Sergentomyia* genus in mountainous foci (Fig. 4, Table 1, Additional file 2: Table S1 (xy sheet)) and a lower frequency is shown in green parts of Fig. 4. VL vectors presence was the most abundant in spatial cluster of Mediterranean region mainly in Raz-o Jargalan and, Maneh-o Samalqan namely Gholaman and Darkesh (Fig. 5, Table 1).

Species diversity

A series of ecological parameters and indices were utilized to characterize the sand fly populations in different areas. Species diversity of sand flies was evaluated using the values of Shannon-wiener index (H'), evenness (E), and richness (S) along with localities of studied areas (Table 3, Additional file 6: Table S4). Analyzing the sand fly species richness, diversity and, evenness indices revealed differences in communities of three bio-climate zones. The highest Shannon diversity index and richness were obtained from the population of *Paraphlebotomus* species with the maximum rate ($H' = 1.098$, $R = 1.224$) in the Mediterranean region, in comparison with other BC areas (Table 3). The second richness index was related to *Sergentomyia* species ($R = 1.154$) in semi-desert community, and the second Shannon diversity index was also related to *Paraphlebotomus* and *Sergentomyia* species with the same rate along with the third most richness index ($H' = 1.039$, $R = 1.060$) in mountainous BC zone (Table 3). *P. papatasi* sand flies indicated no evenness and diversity in all collected BC zones, however its richness was varied from the minimum 0.083 (Paygho) to the maximum 0.936 (Baba Aman) in Mediterranean BC ecotopes. In terms of Jaccard's similarity coefficient for sand fly communities, the similarity of each main locality was ranged from the lowest between Malkesh (mountainous) and Paygho (mountainous), Emarat (semi-arid), Qarahche-Robat (semi-arid), Ashkhane (Mediterranean), Darkesh (Mediterranean) with 25% JSC index similarity to the highest among mountainous localities and, between semi-desert (Emarat, Qarache-Robat) and Mediterranean (Ashkhane, Darkesh) BC zones with 100% JSC index (Table 4, Additional file 6: Table S4).

SP15 and LelF expressions

Quantitative data for the expression profiles of SP15 and LelF of salivary and *Leishmania* genes were summarized in Table 2, Additional file 7: Table S5, in terms of the physiological status (blood-fed, gravid and semi-gravid) and bio-climate classification (Mediterranean, mountainous and semi-arid) for female *P. papatasi* sand flies. Fold change ratios were ranged from the least as 0.12 in June-early to the highest as 24.13 in September-late for SP15 in three Bio-Climature zones (BCz) (Figs. 6, 7, 8). Expression fold of SP15 was ranging from 0.12 in Mediterranean BCz (Gholaman) to 6.41 in semi-arid BCz (Garmeh-Houmeh) in June-early (Fig. 6). One way ANOVA (Kruskal-Wallis test, $KWt=8.548$) of SP15 fold expression in June-early indicated that the medians were significant ($P = 0.0139$, $P < 0.05$) in all three BC zones (Fig. 6, Table 2, Additional file 7: Table S5). Although SP15 fold changes had a marginal increase from Mediterranean BCz to mountainous BCz in June-early, it had also a remarkable rising expression in semi-arid BCz at three times of collection (Fig. 12a, b, c). Mann Whitney *U* test (MWT) was performed to compare SP15 expression fold change ratios between each two BCz in early (June), mid (August), and late (September). Regarding, the obtained data were analyzed and tabulated in Table 2, Additional file 7: Table S5. Expression fold ratios of SP15 were ranged from 0.31 in Mediterranean BCz (Darkesh) to 7.9 in semi-arid BCz (holy tomb of Emamzade-Ebrahim, Sankhast County) in August-mid (Fig. 7, Additional file 4: Table S2). The Kruskal-Wallis test was $KWt=11.19$ for SP15 fold expression in Mid (August) and, the median was significant ($P=0.0037$, $P < 0.05$) in three BC zones (Fig. 12b, Table 2, Additional file 7: Table S5). The expression profiles of SP15 salivary gene were ranged from 0.48 in mountainous BCz (Malkesh) to 24.13 in semi-desert BCz (Kouran district) in September-late (Fig. 8, Additional file 4: Table S2). The Kruskal-Wallis test was $KWt=10.60$ for SP15 fold changes in late (September), and the median was significant ($P=0.0050$, $P < 0.05$) in all BC zones (Fig. 12c, Table 2, Additional file 7: Table S5). The expression profiles of salivary SP15 gene in three regionalized bio-climate locations are as followings: Semi-desert > Mountainous > Mediterranean (Fig. 12a, b, c, Table 2, Additional file 7: Table S5).

The effect of physiological status of blood meal (blood-fed, gravid, and semi-gravid) on expression of LelF gene was examined from natural *Leishmania*-wild-caught *P. papatasi* sand flies in three BC zones. Gravid *P. papatasi* indicated the highest LelF fold expression in Mediterranean BCz (1.9 ± 0.298) (Figs. 11, 12f, Table 2, Additional file 7: Table S5) following semi-gravid in Mediterranean (1.6 ± 0.23) (Figs. 10, 12e, Table 2, Additional file 7: Table S5), and then again gravid with a slight difference in mountainous BC zones (1.55 ± 0.164) (Figs. 10, 12f, Table 2).

Fold changes ratios for LelF expression was ranged from the least as 0.03 for blood-fed status in semi-desert BC zone (Sankhast) (Figs. 9, 12d, Additional file 5: Table S3) to 4.82 as the highest expression fold for gravid status in Mediterranean BC zone (Gholaman) (Figs. 11, 12f, Additional file 5: Table S3). The Kruskal-Wallis test was $KWt=9.046$ for LelF fold expression in gravid status of *P. papatasi*, the median was significant ($P=0.0109$, $P < 0.05$) in three BC zones (Figs. 11, 12f, Table 2, Additional file 7: Table S5). Abdominal position of semi-gravid showed no significant fold changes for LelF expression of *Leishmania* parasites in wild-caught *P. papatasi* in all classified bio climate ecotopes ($KWt=5.285$, $P=0.0712$, $P < 0.05$) (Figs. 10, 12e, Table 2, Additional file 7: Table S5). Statistical analysis for blood-fed status of *P. papatasi* exhibited a significant fold change in LelF expression of *Leishmania* parasites ($KWt = 10.19$, $P = 0.0061$, $P < 0.05$) in three regionalized BC zones (Figs. 9, 12d, Table 2, Additional file 7: Table S5). The expression profiles of LelF from natural *Leishmania*-field-caught *P. papatasi* sand flies were as followings: gravid of A, B and, C > semi-gravid of A, B and, C > blood-fed of A and C in three regionalized BC zones (Fig. 12d, e, f, Table 2, Additional file 7: Table S5).

Discussion

In this investigation, spatial diversity and distribution of leishmaniasis vectors were examined by applying integrative methods (PCA analysis, De Martonne aridity index and, IDW deterministic interpolation in ArcGIS) to regionalize and also to outline the eco-epidemiological risks of VL and CL leishmaniasis in predicted bio-climate foci of northeast of Iran. It was also important to evaluate the fold change expression of salivary protein of field-caught sand fly (SP15) and natural-*Leishmania* sand fly protein (LelF), which has dual characteristics (as a Th1-type adjuvant and as a protection provider) against leishmaniasis. These obtained data are considered as an important foundational step for the establishment of more effective control measures and the surveillance of this vector-borne disease. The criteria were designed in terms of the regionalized BC zones, time of collecting the blood-fed sand flies for comparison of SP15 gene fold change and, different physiological status of sand flies for evaluation of expression level of LelF gene expression.

In line with recent reports in Iran [33], current investigation not only indicated higher diversity for specific species of sand flies in mountainous region only within *Sergentomyia* and *Paraphlebotomus* species, but also has confirmed actually more diversities and richness of vectors of visceral leishmaniasis only in semi-desert and Mediterranean BC communities (Table 3, Additional file 6: Table S4). Vectors of visceral leishmaniasis indicated no considerable diversity in mountainous region except Dartum and Mahnan counties. Of note, in contrast with the recently published data [34], our findings did not show any similarity for the evenness indices in *P. papatasi* population in Northern Khorasan (Table 3). In fact, the evenness were ranged from $E=0.204$ in mountainous BC zone for *Paraphlebotomus* species up to $E=0.721$ in all three BC ecotopes for vectors of visceral leishmaniasis and *Sergentomyia* species (Fig. 4, Table 3, Additional file 6: Table S4). The lack of diversity among *P. papatasi* population and visceral vectors can be caused by the extension of Alborz Mountain range (Aladagh Mountain range) as an important ecological barrier in front of sand fly distribution (Table 3). *P. alexandri* were the most abundant species among *Paraphlebotomus* population, mainly collected from mountainous BC areas at high altitudes (Fig. 3, Table 1, Additional file 1: Figure S1b, Additional file 2: Table S1 (xy sheet)). In fact, in agreement with previous findings in southwest of Iran [5], *P. alexandri* were caught from sylvatic habitats (Qare-Bashloo, Dartum, Gerivan and, Mahnan) near rodents' ecotope and animal shelters in the vicinity of human dwellers (Fig. 3, Additional file 2: Table S1 (xy sheet)). Also, *P. sergenti* was found mostly in rural areas (Paygho, Kohene-Oqaz, Zartanlu) of mountainous BC zone (Fig. 3, Additional file 2: Table S1 (xy sheet)). Our findings were consistent with previous results in Iran [5] which indicated that population of these species tend to be colonized more in places with human inhabitants, and their peridomestic animals even in remote mountainous areas, due to anthropophilic behavior of *P. alexandri* and *P. sergenti*, (Fig. 3, Table 2, Additional file 7: Table S5). One of the important reasons for the higher frequency of visceral leishmaniasis vectors in districts of Gholaman and Darkesh compared to other mountainous BC regions can be caused by the existence and development of dams along with the boundary lines (Fig. 5, Additional file 2: Table S1 (xy sheet)). It is notable that each bio-climate regionalized has its environmental effects on the regulation of SP15 salivary proteins of sand flies associated with gradually increasing levels of SP15 gene expression fold reaching the highest levels of expression late in the season in semi-desert area (Fig. 12 a, b, c), when the environment is drier and sugar sources are scarce (Figs. 6, 7, 8, Table 2, Additional file 7: Table S5) [30]. Also, in addition to blood-feeding of *P. papatasi* on mammalian hosts, feeding with various plants species as a sugar-source diet has various either favorable or adverse effects on life cycle [35, 36] and gonotrophic cycles [37, 38] resulting in increasing and/or decreasing the chance for *Leishmania* transmission (depending on plant species) in irrigated or non-irrigated areas. Gene expression of salivary gland (SP15) in wild-caught *P. papatasi* revealed significant differences between humid western-wet of Mediterranean and semi-arid (A vs C) BC zones, and were also significant between mountainous and semi-arid areas (B vs C) in all times of collection: June (early), August (mid) and, September (late) (Figs. 6, 7, 8, Table 2, Additional file 7: Table S5). Also, there was no significant difference in terms of SP15 fold change in collected blood-fed sand flies between Mediterranean and mountainous BC zones (A vs B) at each time (Table 2, Additional file 7: Table S5). Expression folds of salivary SP15 gene in blood-fed sand flies indicated a significant difference between semi-arid (C) and two other BC regions, however indicated no difference between Mediterranean and mountainous (A vs B) BC areas in June-early, August-mid and, September-late (Table 2, Additional file 7: Table S5). Interestingly, our geospatial of BC analyses indicate that up regulation of SP15 gene expression in semi-arid area confirmed the conceivable correlation of SP15 salivary protein expression with bio-climate variation considering vegetation cover for availability of sugar sources as principle factor [39]. In fact, the expression levels of *P. papatasi* salivary gland genes (SP15) in flies from a dryer habitat (Kouran) (Fig. 8, Additional file 4: Table S2) are higher than the expression levels in those flies from an irrigated area (Darkesh) late in the summer season, when drought may affect the sugar content of plants [38]. Accordingly, correlation between the highest expression of SP44 gene of salivary protein of *P. papatasi* was also confirmed with arid-environment and the rarity of sugar sources in September late [30]. Moreover, aging has been proved as another physiological factor affecting the gene expression profiles; however, it is not considered as significant as genotype or sex determination in insects [40]. Also, similar to what has been observed by other researchers [29], the direct relationship was corroborated between the up-regulation of SP15 transcript and aging.

Expression of LelF gene from *Leishmania* parasites indicated a significant fold in gravid sand flies between semi-desert and two other BC zones (C vs A and B), however indicated no difference between two Mediterranean and Mountainous (A vs B) irrigated areas from early June to late September (Figs. 11, 12f, Table 2, Additional file 7: Table S5). The transcript level of LelF expression from *Leishmania* parasites was found with a significant difference between Mediterranean and semi-desert BC zones (A vs C) for semi-gravid sand flies, while it indicated no fold difference between Mediterranean and mountainous (A vs B), and also, between mountainous and semi-arid (B vs C) BC zones (Figs. 10, 12e, Table 2, Additional file 7: Table S5). Significant difference of LelF fold changes were observed between Mediterranean and semi-arid BC zones (A vs C) in blood-fed *P. papatasi* sand flies, however no difference of gene expression between Mediterranean and mountainous/semi-desert areas was observed (A vs B-C) (Figs. 9, 12d, Table 2, Additional file 7: Table S5). Although a significant difference in LelF gene expression was observed between gravid sand flies of Mediterranean BC zone and blood-fed sand flies of semi-arid region, there was no significant difference between gravid

and semi-gravid sand flies of Mediterranean area with vegetation cover (Figs. 9, 10, 11, Table 2, Additional file 7: Table S5). Nevertheless, testing *Leishmania* gene expression (LelF) in gravid and semi-gravid flies showed also similar results in Mediterranean and mountainous BC regions (Fig. 12e, f, Table 2, Additional file 7: Table S5). However, along with other factors analyzed in this study, gene expression fluctuations for LelF fold change from June (early) to September (late) can be also caused by the existence of some specific plants in three regionalized BC zones (semi-arid, Mediterranean and, mountainous) with the different quality of sugar meals which is attractive for sugar-feeding behavior of *P. papatasi* [38]. Also, expression of genes involved in digestion of blood-meal or sugar meal in blood-fed, semi-gravid, and gravid of sand fly are regulated according to the type of nutrient acquisition [41] which play important roles in *Leishmania* gene expression associated with the distinct environmental BC zones in natural habitats of principle vectors. In contrast with the expression change of salivary protein SP15, the expression fold of natural *Leishmania*-field-caught sand fly gene (LelF) was the lowest in the semi-desert BC region in all three physiological statuses of *P. papatasi* sand flies (gravid, semi-gravid, and, blood-fed) (Fig. 12d, e, f). Level of LelF expression was significantly higher in gravid and semi-gravid sand flies compared to blood-fed sand flies. In fact, in line with recently published data [42]; our research indicated less expression fold of LelF in fresh-blood fed sandflies, and this can be resulted from toxic products of fresh blood-meal digestion. Also, we assumed that according to pruzinova et al. data, as midgut binding is a key factor of *Leishmania* parasites' survival in *P. papatasi*, therefore more availability of sugar sources in irrigated areas (Mediterranean BC zone) can be responsible for higher expression of *Leishmania* promastigote genes and consequently promastigote survival in association with increasing of the salivary protein content [43] to exacerbate the disease in natural geographic habitats.

In the future, it is suggested that the gene regulatory regions of salivary proteins possessing vaccine specificity should be identified against leishmaniasis. Consistent with the results of this study, natural environmental effects which are effectively triggering a significant increase in gene expression fold of specific proteins should be considered with the effects of arid/semi-arid habitat especially for salivary gland protein (SP15). Undoubtedly, the relevance of molecular biology to the mimicry of natural effect of living environment on disease agents is an underlying assumption to assist indoors lab studies using potential applications of gene drive strategy for principle vectors to serve a desirable function of gene at high enough levels, and consequently to control leishmaniasis or vector-borne diseases.

Abbreviations

GIS: Geographic Information System, LelF: *Leishmania* eukaryotic initiation factor, SP: Salivary protein, *SaLeish*: Saliva+*Leishmania* proteins, DTH: delayed-type hypersensitivity, DC: dendritic cells, BCz: Bio Climate zone, CL: Cutaneous leishmaniasis, VL: Visceral leishmaniasis, PCA: Principal Component Analysis, CIM: Clustering Integration Method, IDW: Inverse Distance Weighting, JSC: Jaccard's similarity coefficient.

Declarations

Acknowledgements

This paper has been extracted from the Ph.D thesis written by Ali Bordbar, Molecular Systematics Laboratory, Department of Medical Parasitology, at Pasteur Institute of Iran. The authors are very grateful to Ms. Sahar Ebrahimi for helping with the field work, and her assistance in sand fly dissection, morphological identifications, and laboratory experiments. We are gratefully thanked to Mr. Gholamreza Ebrahimi for helping with the field work that generously provided accommodation and transfer facilities. We appreciate indigenous people in rural areas of Northern Khorasan province that sincerely allowed us to collect sand flies from interior properties, stables and from their agricultural farms.

Authors' Contributions

AB and PP conceived and designed the idea of the research. AB performed Data curation, Conceptualization, Investigation, Methodology, Software, Visualization, original draft, Writing–review & editing. PP performed Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Supervision, Validation, Visualization, Writing–review & editing.

Funding

Funding was available through the studentship granted from Pasteur Institute of Iran to Ali Bordbar, Ph.D. student and a research fund granted by the National Institute for Medical Research Development (NIMAD) (Project No. 973166) awarded to Prof. P. Parvizi. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

All essential data generated or analyzed during this study are presented in this published article and supplementary section. The raw datasets are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author details

Molecular Systematics Laboratory, Parasitology Department, Pasteur Institute of Iran, Tehran, Iran.

References

1. Berman JD. Human leishmaniasis: clinical, diagnostic, and chemotherapeutic developments in the last 10 years. *Clinical infectious diseases*. 1997 Apr 1;24(4):684-703.
2. World Health Organization, *Leishmaniasis*, Fact sheet, 14 March 2019;<https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>.
3. Alvar J, Vélez ID, Bern C, Herrero M, Desjeux P, Cano J, Jannin J, den Boer M, WHO Leishmaniasis Control Team. Leishmaniasis worldwide and global estimates of its incidence. *PloS one*. 2012 May 31;7(5):e35671.
4. Gavvani AS, Mohite H, Edrissian GH, Mohebalı M, Davies CR. Domestic dog ownership in Iran is a risk factor for human infection with *Leishmania infantum*. *The American journal of tropical medicine and hygiene*. 2002 Nov 1;67(5):511-5.
5. Ebrahimi S, Bordbar A, Rastaghi AR, Parvizi P. Spatial distribution of sand fly species (Psychodidae: Phlebotominae), ecological niche, and climatic regionalization in zoonotic foci of cutaneous leishmaniasis, southwest of Iran. *Journal of Vector Ecology*. 2016 Jun;41(1):103-9.
6. Javadian E, Nadim A, Tahvildare-Bidruni GH, Assefi V. Epidemiology of cutaneous leishmaniasis in Iran: B. Khorassan Part V: Report on a focus of zoonotic cutaneous leishmaniasis in Esferayen. *Bulletin de la Societe de Pathologie Exotique et de ses Filiales*. 1976;69(2):140-3.
7. Bhunia GS, Shit PK. Exploring Ecology and Associated Disease Pattern. In *Geospatial Analysis of Public Health 2019* (pp. 139-198). Springer, Cham.
8. Ebrahimi S, Bordbar A, Parvizi P. Genetic dynamics in the sand fly (Diptera: Psychodidae) nuclear and mitochondrial genotypes: evidence for vector adaptation at the border of Iran with Iraq. *Parasites & vectors*. 2016 Dec;9(1):319.
9. Zhang Q, Hou X, Li FY, Niu J, Zhou Y, Ding Y, Zhao L, Li X, Ma W, Kang S. Alpha, beta and gamma diversity differ in response to precipitation in the Inner Mongolia grassland. *PLoS One*. 2014 Mar 27;9(3):e93518.
10. Lafferty KD. The ecology of climate change and infectious diseases. *Ecology*. 2009 Apr;90(4):888-900.
11. Schowalter TD, Noriega JA, Tscharntke T. Insect effects on ecosystem services, 2018—Introduction.
12. Campbell-Lendrum D, Manga L, Bagayoko M, Sommerfeld J. Climate change and vector-borne diseases: what are the implications for public health research and policy?. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2015 Apr 5;370(1665):20130552.
13. Desjeux P. Leishmaniasis: current situation and new perspectives. *Comparative immunology, microbiology and infectious diseases*. 2004 Sep 1;27(5):305-18.
14. Gillespie PM, Beaumier CM, Strych U, Hayward T, Hotez PJ, Bottazzi ME. Status of vaccine research and development of vaccines for leishmaniasis. *Vaccine*. 2016 Jun 3;34(26):2992-5.
15. Bordbar A, Bagheri KP, Ebrahimi S, Parvizi P. Bioinformatics analyses of immunogenic T-cell epitopes of LelF and PpSP15 proteins from *Leishmania major* and sand fly saliva used as model antigens for the design of a multi-epitope vaccine to control leishmaniasis. *Infection, Genetics and Evolution*. 2020 Jan 10:104189.
16. Oliveira F, Lawyer PG, Kamhawi S, Valenzuela JG. Immunity to distinct sand fly salivary proteins primes the anti-*Leishmania* immune response towards protection or exacerbation of disease. *PLoS neglected tropical diseases*. 2008 Apr 16;2(4):e226.
17. Valenzuela JG, Belkaid Y, Garfield MK, Mendez S, Kamhawi S, Rowton ED, Sacks DL, Ribeiro JM. Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. *Journal of Experimental Medicine*. 2001 Aug 6;194(3):331-42.
18. Belkaid Y, Valenzuela JG, Kamhawi S, Rowton E, Sacks DL, Ribeiro JM. Delayed-type hypersensitivity to *Phlebotomus papatasi* sand fly bite: An adaptive response induced by the fly?. *Proceedings of the National Academy of Sciences*. 2000 Jun 6;97(12):6704-9.
19. Oliveira F, Rowton E, Aslan H, Gomes R, Castrovinci PA, Alvarenga PH, Abdeladhim M, Teixeira C, Meneses C, Kleeman LT, Guimarães-Costa AB. A sand fly salivary protein vaccine shows efficacy against vector-transmitted cutaneous leishmaniasis in nonhuman primates. *Science translational medicine*. 2015 Jun 3;7(290):290ra90-.
20. Gomes R, Oliveira F. The immune response to sand fly salivary proteins and its influence on *Leishmania* *Frontiers in immunology*. 2012 May 11;3:110.
21. Skeiky YA, Kennedy M, Kaufman D, Borges MM, Guderian JA, Scholler JK, Owendale PJ, Picha KS, Morrissey PJ, Grabstein KH, Campos-Neto A. LelF: a recombinant *Leishmania* protein that induces an IL-12-mediated Th1 cytokine profile. *The Journal of Immunology*. 1998 Dec 1;161(11):6171-9.

22. Koutsoni O, Barhoumi M, Guizani I, Dotsika E. *Leishmania* eukaryotic initiation factor (LeIF) inhibits parasite growth in murine macrophages. PLoS One. 2014 May 15;9(5):e97319.
23. Coler RN, Goto Y, Bogatzki L, Raman V, Reed SG. Leish-111f, a recombinant polyprotein vaccine that protects against visceral Leishmaniasis by elicitation of CD4+ T cells. Infection and immunity. 2007 Sep 1;75(9):4648-54.
24. Hodgins-Davis A, Townsend JP. Evolving gene expression: from G to E to G× E. Trends in Ecology & Evolution. 2009 Dec 1;24(12):649-58.
25. Setianto A, Triandini T. Comparison of kriging and inverse distance weighted (IDW) interpolation methods in lineament extraction and analysis. Journal of Applied Geology. 2013;5(1).
26. Magurran AE. Measuring biological diversity. John Wiley & Sons; 2013 Apr 18.
27. Nadim A, Javadian E. Key for the species identification of sand flies (Diptera phlebotominae) of Iran. Iranian J Publ Hlth. 1976;5(1):25-8.
28. Lewis DJ. A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae). 1982 Bull Br Mus Nat Hist Entomol Ser 45:121–209.
29. Hosseini-Vasoukolaei N, Idali F, Khamesipour A, Yaghoobi-Ershadi MR, Kamhawi S, Valenzuela JG, Edalatkhah H, Arandian MH, Mirhendi H, Emami S, Jafari R. Differential expression profiles of the salivary proteins SP15 and SP44 from *Phlebotomus papatasi*. Parasites & vectors. 2016 Dec;9(1):357.
30. Coutinho-Abreu IV, Mukbel R, Hanafi HA, Fawaz EY, El-Hossary SS, Wadsworth M, Stayback G, Pitts DA, Abo-Shehada M, Hoel DF, Kamhawi S. Expression plasticity of *Phlebotomus papatasi* salivary gland genes in distinct ecotopes through the sand fly season. BMC ecology. 2011 Dec;11(1):24.
31. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. methods. 2001 Dec 1;25(4):402-8.
32. Hamilton AJ. Species diversity or biodiversity?. Journal of Environmental Management. 2005 Apr 1;75(1):89-92.
33. Abedi-Astaneh F, Akhavan AA, Shirzadi MR, Rassi Y, Yaghoobi-Ershadi MR, Hanafi-Bojd AA, Akbarzadeh K, Nafar-Shalamzari R, Parsi S, Abbasi A, Raufi H. Species diversity of sand flies and ecological niche model of *Phlebotomus papatasi* in central Iran. Acta tropica. 2015 Sep 1;149:246-53.
34. Arzamani K, Vatandoost H, Rassi Y, Akhavan AA, Abai MR, Alavinia M, Akbarzadeh K, Mohebbali M, Rafizadeh S. Richness and Diversity of Phlebotomine Sand Flies (Diptera: Psychodidae) in North Khorasan Province, Northeast of Iran. Journal of arthropod-borne diseases. 2018 Sep;12(3):232.
35. Schlein Y, Jacobson RL. Sugar meals and longevity of the sandfly *Phlebotomus papatasi* in an arid focus of *Leishmania major* in the Jordan Valley. Medical and Veterinary Entomology. 1999 Mar;13(1):65-71.
36. Yuval B. Populations of *Phlebotomus papatasi* (Diptera: Psychodidae) and the risk of *Leishmania major* transmission in three Jordan Valley habitats. Journal of medical entomology. 1991 Jul 1;28(4):492-5.
37. Schlein Y, Jacobson RL. Mortality of *Leishmania major* in *Phlebotomus papatasi* caused by plant feeding of the sand flies. The American journal of tropical medicine and hygiene. 1994 Jan 1;50(1):20-7.
38. Schlein Y, Jacobson RL. Photosynthesis modulates the plant feeding of *Phlebotomus papatasi* (Diptera: Psychodidae). Journal of medical entomology. 2000 May 1;37(3):319-24.
39. Schlein Y, Jacobson RL. Linkage between susceptibility of *Phlebotomus papatasi* to *Leishmania major* and hunger tolerance. Parasitology. 2002 Oct;125(4):343-8.
40. Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nature genetics. 2001 Nov 26;29(4):389.
41. Ramalho-Ortigão JM, Kamhawi S, Joshi MB, Reynoso D, Lawyer PG, Dwyer DM, Sacks DL, Valenzuela JG. Characterization of a blood activated chitinolytic system in the midgut of the sand fly vectors *Lutzomyia longipalpis* and *Phlebotomus papatasi*. Insect molecular biology. 2005 Dec;14(6):703-12.
42. Pruzinova K, Sadlova J, Myskova J, Lestinova T, Janda J, Volf P. *Leishmania* mortality in sand fly blood meal is not species-specific and does not result from direct effect of proteinases. Parasites & vectors. 2018 Dec;11(1):37.
43. Prates DB, Santos LD, Miranda JC, Souza AP, Palma MS, Barral-netto M, Barral A. Changes in amounts of total salivary gland proteins of *Lutzomyia longipalpis* (Diptera: Psychodidae) according to age and diet. Journal of medical entomology. 2014 Oct 29;45(3):409-13.

Tables

Table 1 Relative abundance and distribution of sandflies (isolated by gender) collected from main different locations of Northern Khorasan province.

F: female, M: male, Codes: classified for R analysis.

Table 2 Comparison of salivary gland and *Leishmania* gene expression levels (SP15 & LeIF) isolated from field-caught *P. papatasi* in three Bio-climate zones of Northern Khorasan province.

Climate regionalization		Mountainous														Semi-Desert				Mediterranean				Total (%)													
Main District		Bojnord				Esfarayen				Shirvan				Jajarm				Maneho-Samalqan		Razo-Jargalan																	
Focal Locations		Baba Aman	Malkesh	Ghareh Bashloo	Qarah Khan Bandi	Quch Qaleh-Olya	Paygho	Dartum	Gerivan	Mahnan	Chalu	Kohneh Ogaz		Zartanlu	Emarat	Qarah Cheh Robat	Ashkhane	Darkesh	Gholaman																		
Genus	Code	A		B		C		D		E		F		G		H		I		J		K		L		M		N		O		P		Q			
	Species	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M				
<i>Phlebotomus</i>	<i>P. papatasi</i>	100	14	0	8	14	6	86	2	88	2	98	46	14	18	44	6	16	6	16	7	24	10	19	4	32	7	14	3	41	5	15	2	29	4	800 (68.55)	
	<i>P. kandelaki</i>	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	4	0	10 (0.86)
	<i>P. major group</i>	0	0	0	0	0	0	0	0	0	0	4	0	2	0	0	0	2	0	0	0	0	0	0	0	2	0	2	1	2	0	5	0	3	0	23 (1.97)	
	<i>P. chinensis group</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0	5	0	2	0	6	1	3	0	21 (1.8)	
	<i>P. alexandri</i>	12	0	0	0	20	2	0	0	0	0	14	2	12	12	22	2	22	2	2	0	17	1	9	0	0	0	0	0	2	0	0	0	0	0	0	153 (13.11)
	<i>P. ansari</i>	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2 (0.17)
	<i>P. caucasicus</i>	0	0	0	0	4	2	0	0	0	0	4	0	0	0	0	0	2	0	0	2	0	5	0	4	0	0	0	0	0	0	2	0	0	0	0	25 (2.14)
	<i>P. sergenti</i>	0	0	0	0	0	0	2	0	0	0	0	6	0	0	0	0	0	0	0	0	4	1	8	1	4	1	0	3	0	0	2	1	6	0	0	39 (3.34)
<i>Sergentomyia</i>	<i>S. baghdadis</i>	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	17 (1.46)
	<i>S. sumbarica</i>	4	0	0	0	0	0	4	0	0	0	0	0	4	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14 (1.2)
	<i>S. dentata</i>	2	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	10 (0.86)	
	<i>S. sintoni</i>	2	0	0	0	0	2	0	0	14	0	0	0	0	4	0	0	0	0	0	7	0	2	1	3	0	1	0	6	0	3	0	0	1	7	53 (4.54)	
Total (%)		120	14	0	8	54	12	96	2	102	2	120	54	8	34	66	8	48	8	18	20	42	26	30	15	37	10	27	10	47	12	29	11	40	11	1167	
		134	8	66	8	54	12	98	104	174	68	74	56	38	68	45	47	37	59	40	51	45	37	47	37	59	40	51	45	37	47	37	59	40	51	45	
		(11.48)	(0.68)	(5.65)	(0.68)	(8.4)	(8.91)	(8.4)	(8.91)	(14.91)	(5.82)	(6.34)	(4.8)	(3.26)	(5.82)	(3.86)	(4.02)	(3.17)	(5)	(3.42)	(4.37)	(3.26)	(5.82)	(4.02)	(3.17)	(5)	(3.42)	(4.37)	(3.26)	(5.82)	(4.02)	(3.17)	(5)	(3.42)	(4.37)		

* Bojnord, Shirvan, Garmeh, Faruj, and Esfarayen Counties.

** Raz-o Jargalan, and Maneh-o Samalqan Counties.

• Jajarm, southern parts of Garmeh and, western areas of Esfarayen Counties, SEM: Standard Error of Mean, ^{MW}, Mann-Whitney U Test

{ The remaining 7 variables are shown in Supplementary 3, S: Significant, N: No significant

Table 3 Biodiversity analysis for different species of sand flies in bio-climate zones of Northern Khorasan province.

Isolated source	Status	Climate ecotopes			Frequency of Distribution of Gene expression			P value ($P < 0.05$) ^{MW}				
		(A)	(B)	(C)	Mean \pm SEM of All 14 sites Expression Fold			All 14 sites Expression Fold				
		Mediterranean**	Mountainous*	Semi arid•	(A)	(B)	(C)	A vs B	A vs C	B vs C		
<i>P. papatasi</i>	Salivary protein SP15 (Fold Change)	Blood Fed 7 out of 14 [†]	Early (June)	1.20	0.32	1.67	1.40429 \pm 0.382118	1.585 \pm 0.419459	2.77571 \pm 0.434103	0.6295 ^N	0.0101 ^S	0.0180 ^S
				0.96	0.54	3.14						
				0.42	0.49	2.23						
				1.70	1.94	2.12						
				0.20	1.73	2.85						
				0.56	2.92	5.63						
			2.80	3.89	3.88							
			Mid (Aug) 7 out of 14 [†]	1.21	1.58	1.99	1.87071 \pm 0.423889	2.095 \pm 473827	4.13571 \pm 0.575062	0.7652 ^N	0.0035 ^S	0.0054 ^S
			0.31	0.29	3.88							
			0.65	0.57	2.81							
			1.77	2.10	2.70							
			3.23	3.42	3.50							
			0.53	0.85	6.31							
			3.83	4.40	7.37							
			Late (Sep) 7 out of 14 [†]	1.79	1.92	2.30	3.91929 \pm 1.00625	3.37286 \pm 0.770445	6.70214 \pm 1.44592	0.5972 ^N	0.0115 ^S	0.0030 ^S
			0.55	0.48	4.37							
			1.19	1.60	3.62							
			3.83	3.72	4.41							
			1.97	1.48	4.10							
			3.27	1.39	9.79							
			4.80	6.49	24.13							
<i>Leishmania parasite (P. papatasi)</i>	LeIF (Fold Change)	Gravid 7 out of 14 [†]		0.770	1.090	0.920	1.90557 \pm 0.298908	1.55971 \pm 0.164003	0.997929 \pm 108543	0.5502 ^N	0.0101 ^S	0.0123 ^S
				2.070	1.500	1.340						
				0.920	1.700	0.590						
				1.810	1.210	0.430						
				0.993	0.940	0.790						
				1.440	1.110	0.970						
			2.130	1.830	1.120							
			Semi-Gravid 7 out of 14 [†]	0.460	0.820	0.410	1.63629 \pm 0.230047	1.29536 \pm 0.17331	0.956429 \pm 159147	0.3012 ^N	0.0326 ^S	0.1543 ^N
			1.530	1.720	1.360							
			0.670	0.560	0.600							
			0.980	0.850	0.530							
			2.070	1.640	0.840							
			2.910	0.990	1.470							
			2.414	2.210	1.320							
			Blood-Fed 7 out of 14 [†]	0.210	0.16	0.11	0.577857 \pm 0.119376	1.07714 \pm 0.188427	0.355 \pm 0.0988436	0.0596 ^N	0.0886 ^N	0.0035 ^S
			1.190	1.90	0.17							
			0.090	1.70	1.02							
			1.280	0.81	0.18							
	0.190	0.24	0.09									
	0.340	0.11	0.11									
	1.130	0.39	0.52									

Main Regionalized Zones	Location	Code	<i>P. papatasi</i>			Vectors of Visceral Leishmaniasis (VL)			<i>Paraphlebotomus</i> species			<i>Sergentomyia</i> species		
			Richness (R)	Diversity (H _W)	Evenness (E)	Richness (R)	Diversity (H _W)	Evenness (E)	Richness (R)	Diversity (H _W)	Evenness (E)	Richness (R)	Diversity (H _W)	Evenness (E)
Mountainous	Baba Aman	A	0.936	0.000	0.000	0.000	0.000	0.000	0.288	0.000	0.000	1.060	1.039	0.568
	Malkesh	B	0.353	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Ghareh	C	0.223	0.000	0.000	0.000	0.000	0.000	0.547	0.729	0.380	0.500	0.367	0.315
	Bashloo													
	Qarah Khan Bandi	D	0.106	0.000	0.000	0.000	0.000	0.000	0.707	0.000	0.000	0.707	0.693	0.721
	Quch Qaleh-Olya	E	0.105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.267	0.000	0.000
	Paygho	F	0.083	0.000	0.000	0.500	0.000	0.000	0.588	0.925	0.495	0.000	0.000	0.000
	Dartum	G	0.176	0.000	0.000	1.000	0.693	0.721	0.204	0.000	0.000	0.707	0.693	0.721
	Gerivan	H	0.141	0.000	0.000	0.000	0.000	0.000	0.204	0.000	0.000	0.000	0.000	0.000
	Mahnán	I	0.213	0.000	0.000	1.000	0.693	0.721	0.392	0.271	0.204	1.000	0.693	0.721
	Chalu	J	0.208	0.000	0.000	0.000	0.000	0.000	1.060	1.039	0.568	0.377	0.000	0.000
	Kohneh Ogaz	K	0.171	0.000	0.000	0.000	0.000	0.000	0.530	0.970	0.528	0.707	0.000	0.000
Zartanlı	L	0.208	0.000	0.000	0.000	0.000	0.000	0.707	1.036	0.567	0.500	0.000	0.000	
Semi-desert	Emarat	M	0.160	0.000	0.000	1.000	0.693	0.721	1.000	0.000	0.000	1.154	0.636	0.641
	Qarah Cheh Robot	N	0.242	0.000	0.000	0.948	1.029	0.564	0.577	0.000	0.000	0.755	0.410	0.353
Mediterranean	Ashkhane	O	0.147	0.000	0.000	1.000	0.693	0.721	1.224	1.098	0.606	0.577	0.000	0.000
	Darkesh	P	0.242	0.000	0.000	0.801	0.922	0.548	0.377	0.000	0.000	0.707	0.000	0.000
	Gholaman	Q	0.174	0.000	0.000	0.948	1.089	0.600	0.000	0.000	0.000	0.353	0.000	0.000

H_W = Shannon-Weiner index, R = Menhinick Richness index, E = Shannon Evenness index

Table 4 Jaccard's similarity coefficient calculated for all main localities of collecting sites of sand flies in north east of Iran.

BC* zones	Locality code**	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Mountainous	A	1	0.33	1	1	0.67	0.75	0.75	0.67	0.75	1	1	1	0.75	0.75	0.75	0.75	0.50
	B		1	0.33	0.33	0.50	0.25	0.25	0.50	0.25	0.33	0.33	0.33	0.25	0.25	0.25	0.25	0.33
	C			1	1	0.67	0.75	0.75	0.67	0.75	1	1	1	0.75	0.75	0.75	0.75	0.50
	D				1	0.67	0.75	0.75	0.67	0.75	1	1	1	0.75	0.75	0.75	0.75	0.50
	E					1	0.50	0.50	0.33	0.50	0.67	0.67	0.67	0.50	0.50	0.50	0.50	0.67
	F						1	1	0.50	1	0.75	0.75	0.75	1	1	1	1	0.75
	G							1	0.50	1	0.75	0.75	0.75	1	1	1	1	0.75
	H								1	0.50	0.67	0.67	0.67	0.50	0.50	0.50	0.50	0.25
	I									1	0.75	0.75	0.75	1	1	1	1	0.75
	J										1	1	1	0.75	0.75	0.75	0.75	0.50
	K											1	1	0.75	0.75	0.75	0.75	0.50
	L												1	0.75	0.75	0.75	0.75	0.50
Semi-arid	M													1	1	1	1	0.75
	N														1	1	1	0.75
Mediterranean	O															1	1	0.75
	P																1	0.75
	Q																	1

* Bio-Climat, ** as shown in Table 1.

Additional File Legends

Additional file 1: Figure S1. Northern Khorasan province was incriminated based on significant factors a) Height, b) Annual rainfall, c) Temperature, d) Humidity using deterministic spatial interpolation method in GIS.

Additional file 2: Table S1. Distribution of various sand fly species based on species diversity collected from three climatic zones of Northern Khorasan province, north east of Iran (2017-2018).

Additional file 3: Figure S2. Bioclimatic regionalization of Northern Khorasan province generated by the IDW method in ArcGIS.

Additional file 4: Table S2. SP15 BloodFed-Fold Expression of wild-caught *P. papatasi* in three bio climate zones at three different times.

Additional file 5: Table S3. LeIF Fold Expression of wild-caught *P. papatasi* in Gravid, SemiGravid, and BloodFed statuses with respect to three bio climate zones.

Additional file 6: Table S4. Analyses of Jaccard's similarity index and sand fly species diversity based on Shannon-Wiener index, species richness and evenness index.

Additional file 7: Table S5. Analyses of 7 remaining samples related to comparison of salivary gland and *Leishmania* gene expression levels (SP15 & LeIF) isolated from field-caught *P. papatasi* in three Bio-climate zones of Northern Khorasan province.

Figures

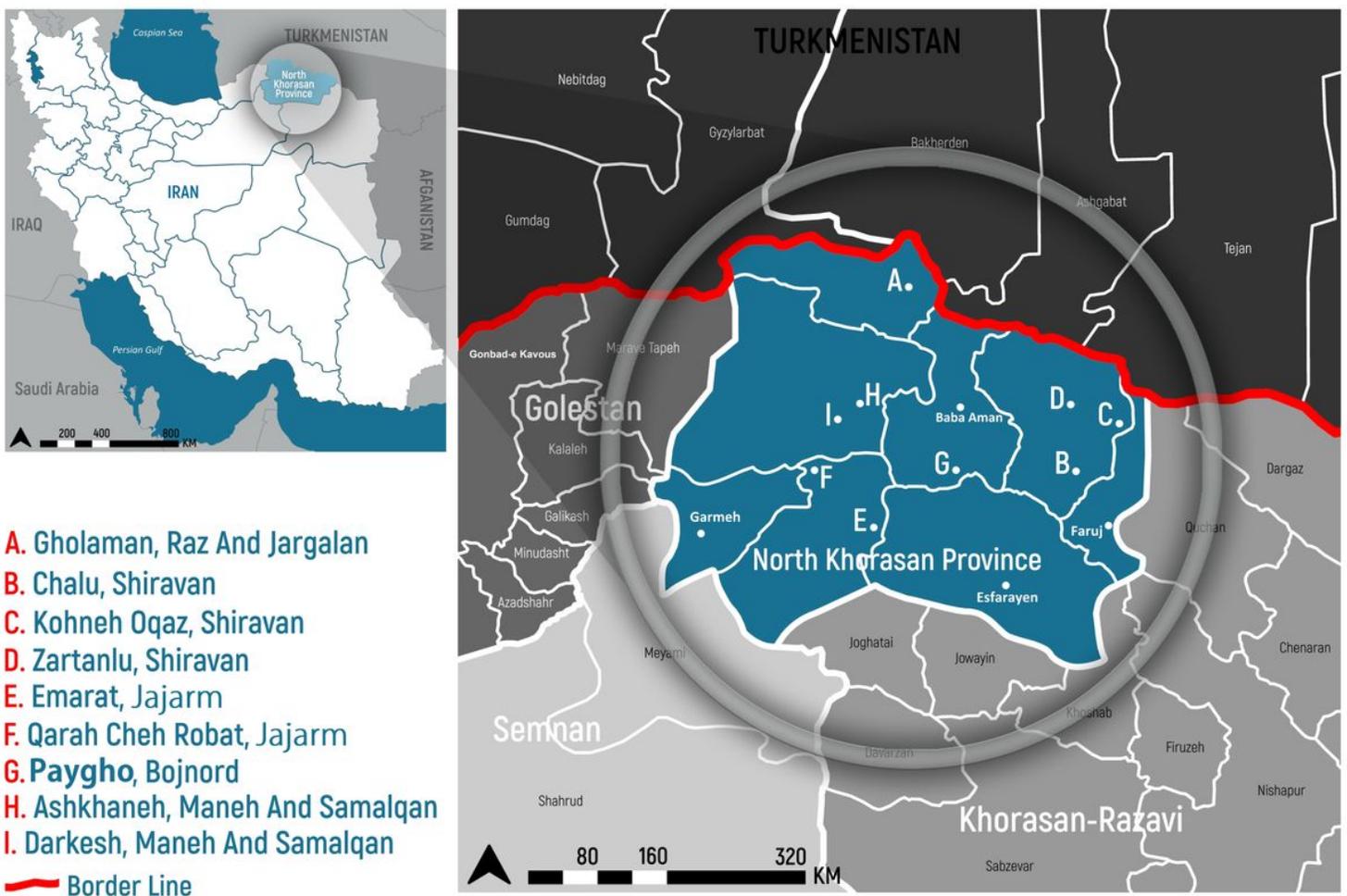


Figure 1

Geographical location of visceral and cutaneous leishmaniasis in north east of Iran. Main collected sites of endemic areas are marked with letters.

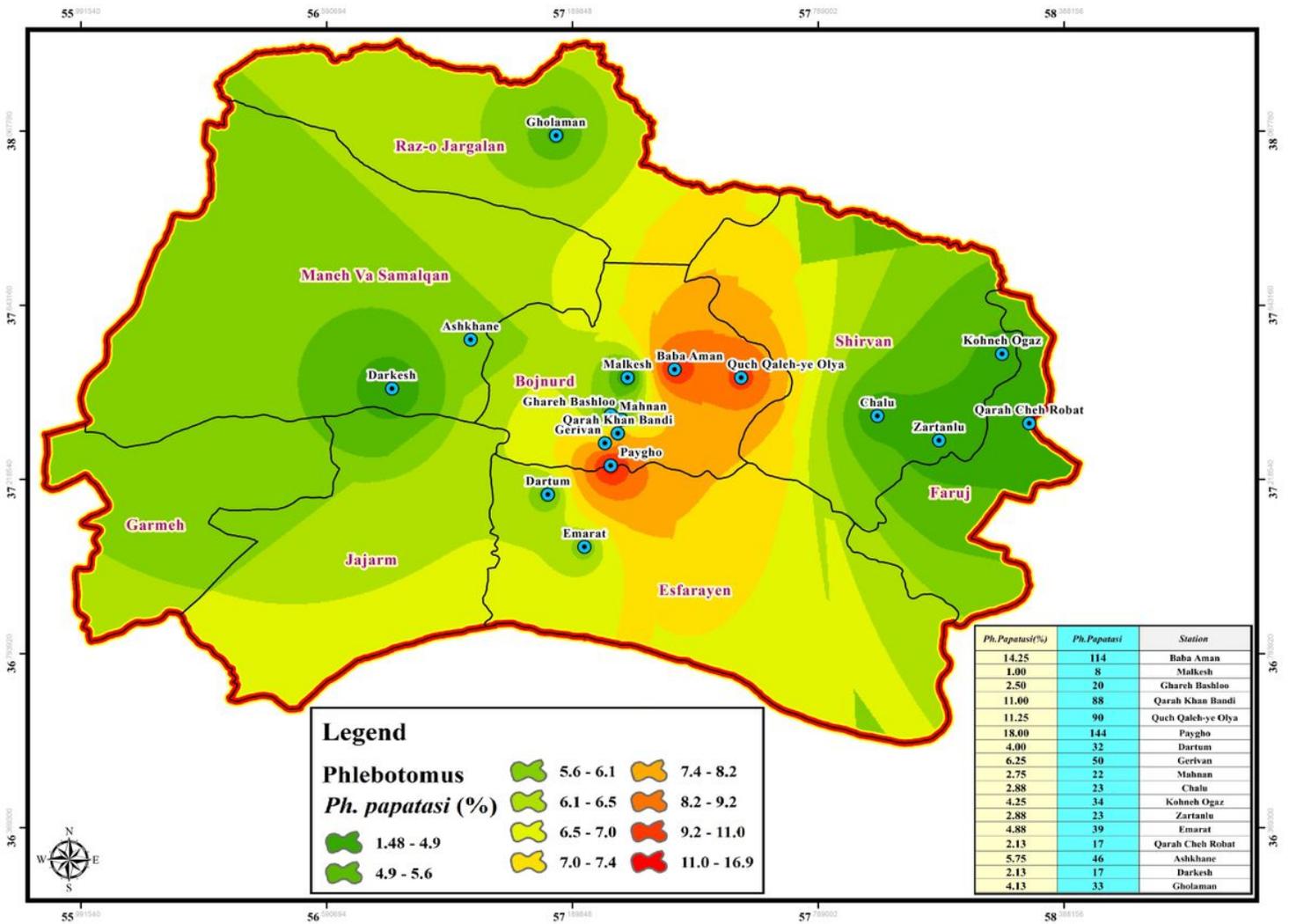


Figure 2

Distribution of principle vector of CL, Phlebotomus papatasi sand fly as hazardous map in north of Iran using IDW interpolation method in ArcGIS.

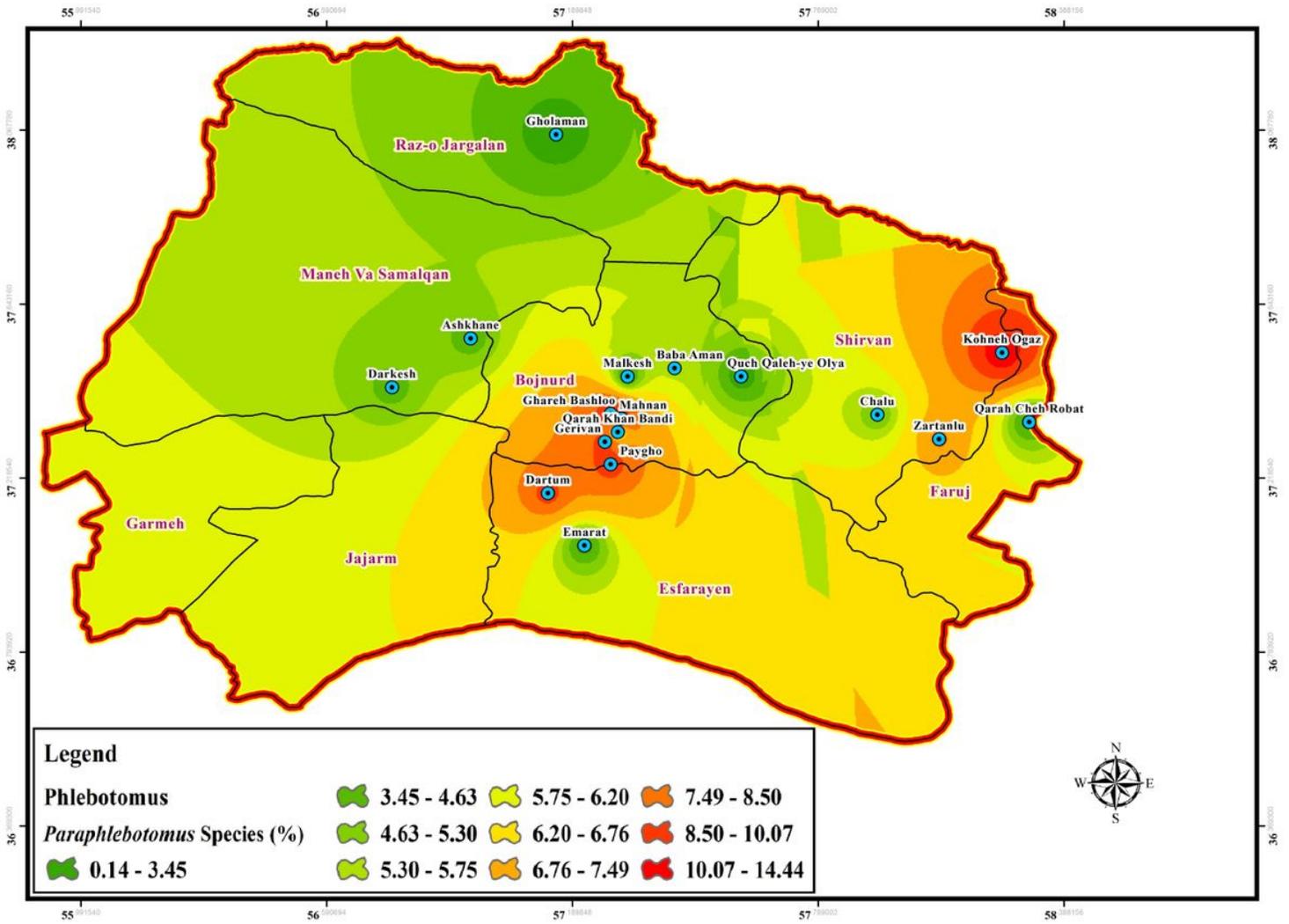


Figure 3

Hazardous map of Paraphlebotomus species, principle vectors of CL/or potential vectors of VL in north of Iran using ArcGIS.

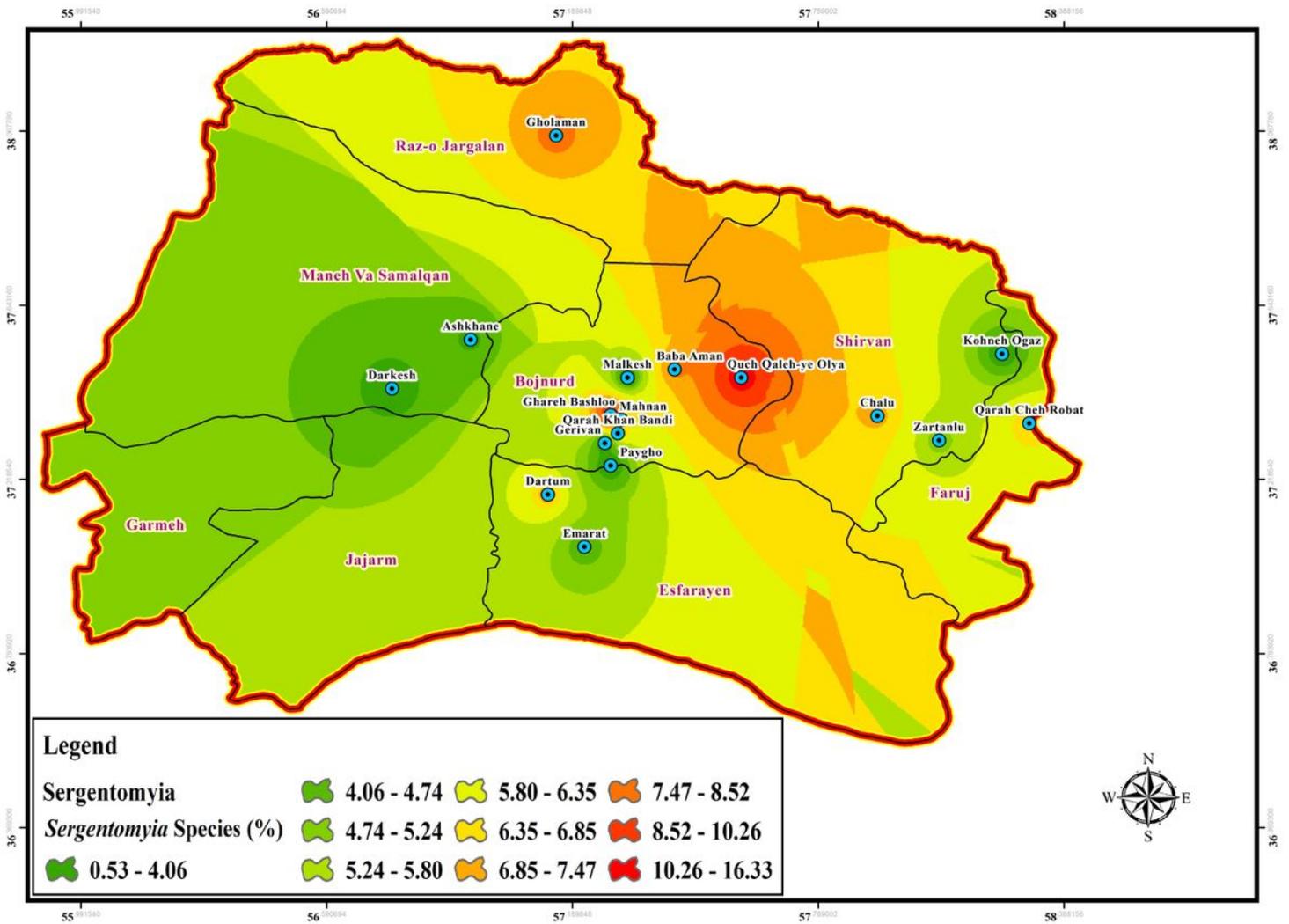


Figure 4

Spatial distribution of *Sergentomyia* species in northern Khorasan, Iran.

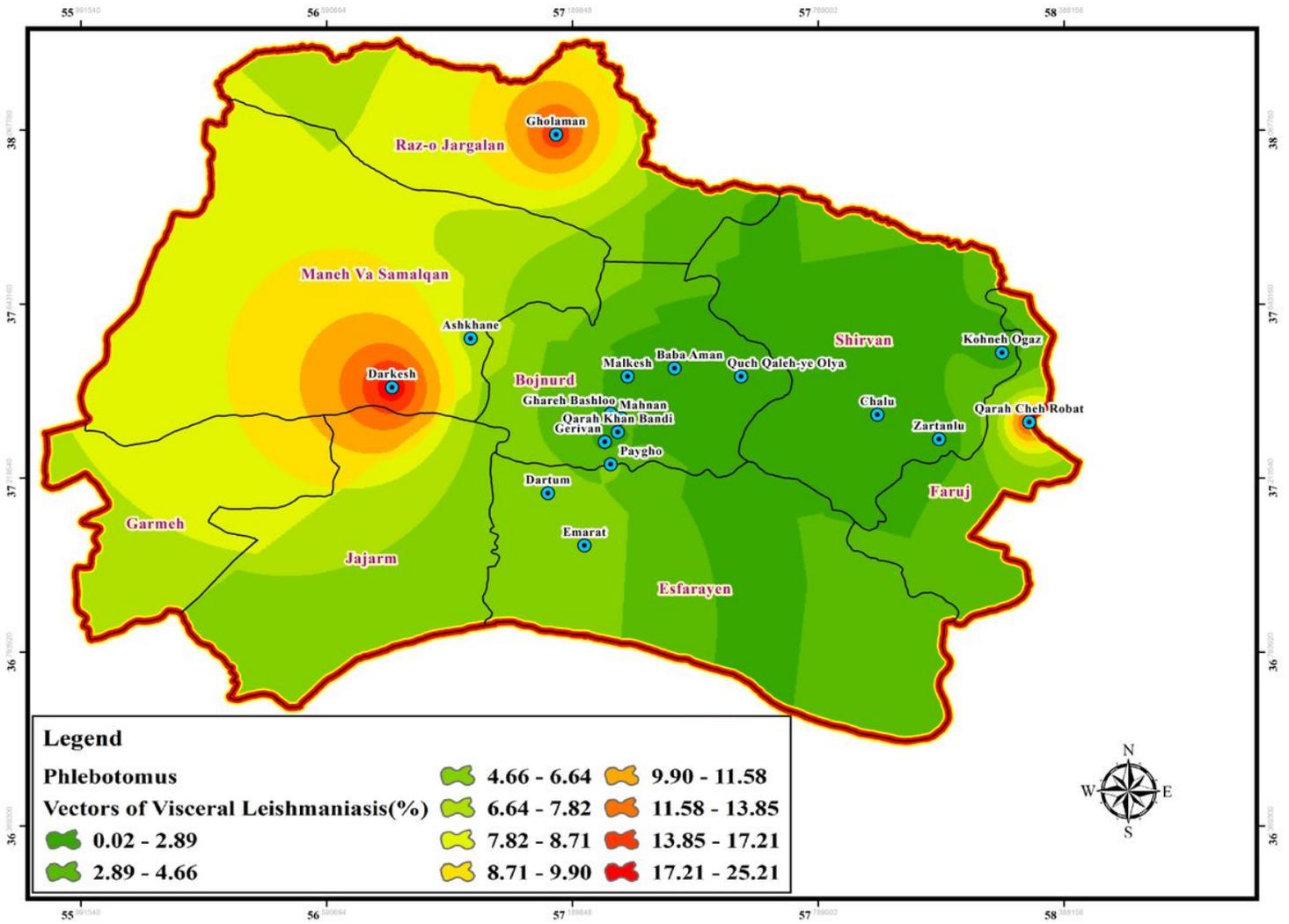


Figure 5

Dispersion of principle vectors of visceral leishmaniasis in north of Iran.

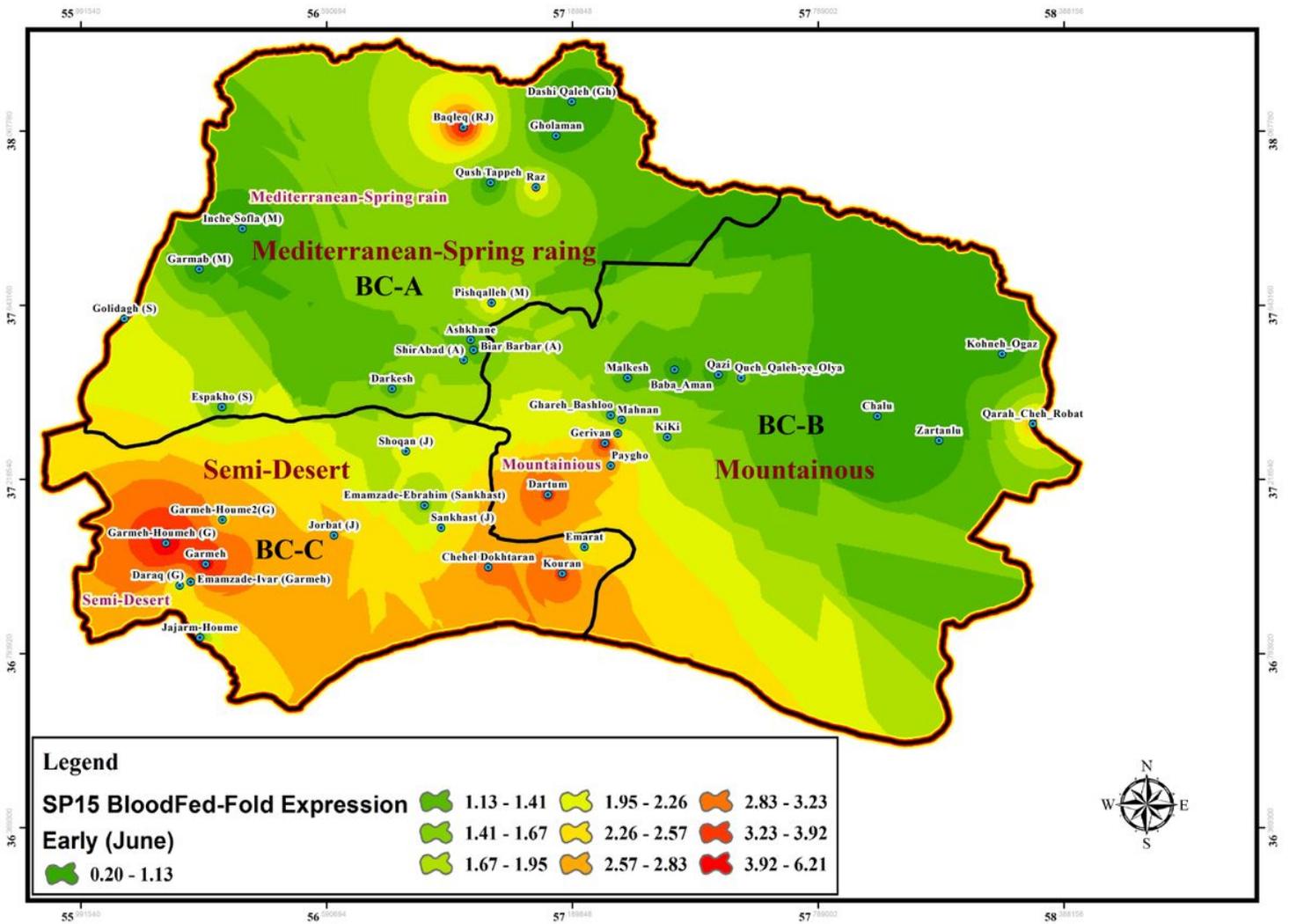


Figure 6

Gene expression changes of SP15 salivary protein in wild-collected *Ph. papatasi* sand flies based on bio-climate zones and time of collection, June (Early), BC: Bio Climate zones.

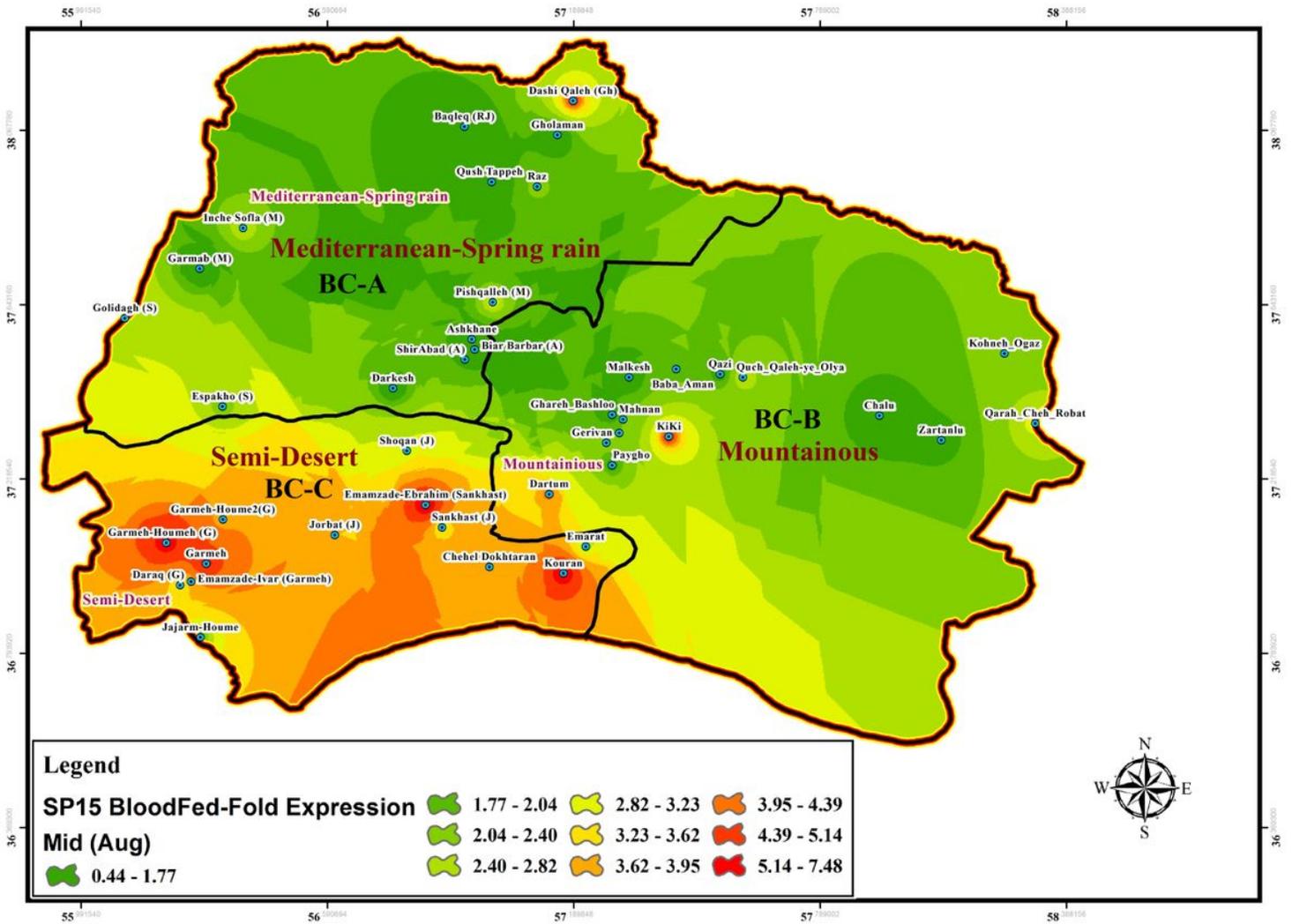


Figure 7
 Gene expression changes of SP15 salivary protein in wild-collected *Ph. papatasi* sand flies based on bio-climate zones and time of collection, August (Mid), BC: Bio Climate areas.

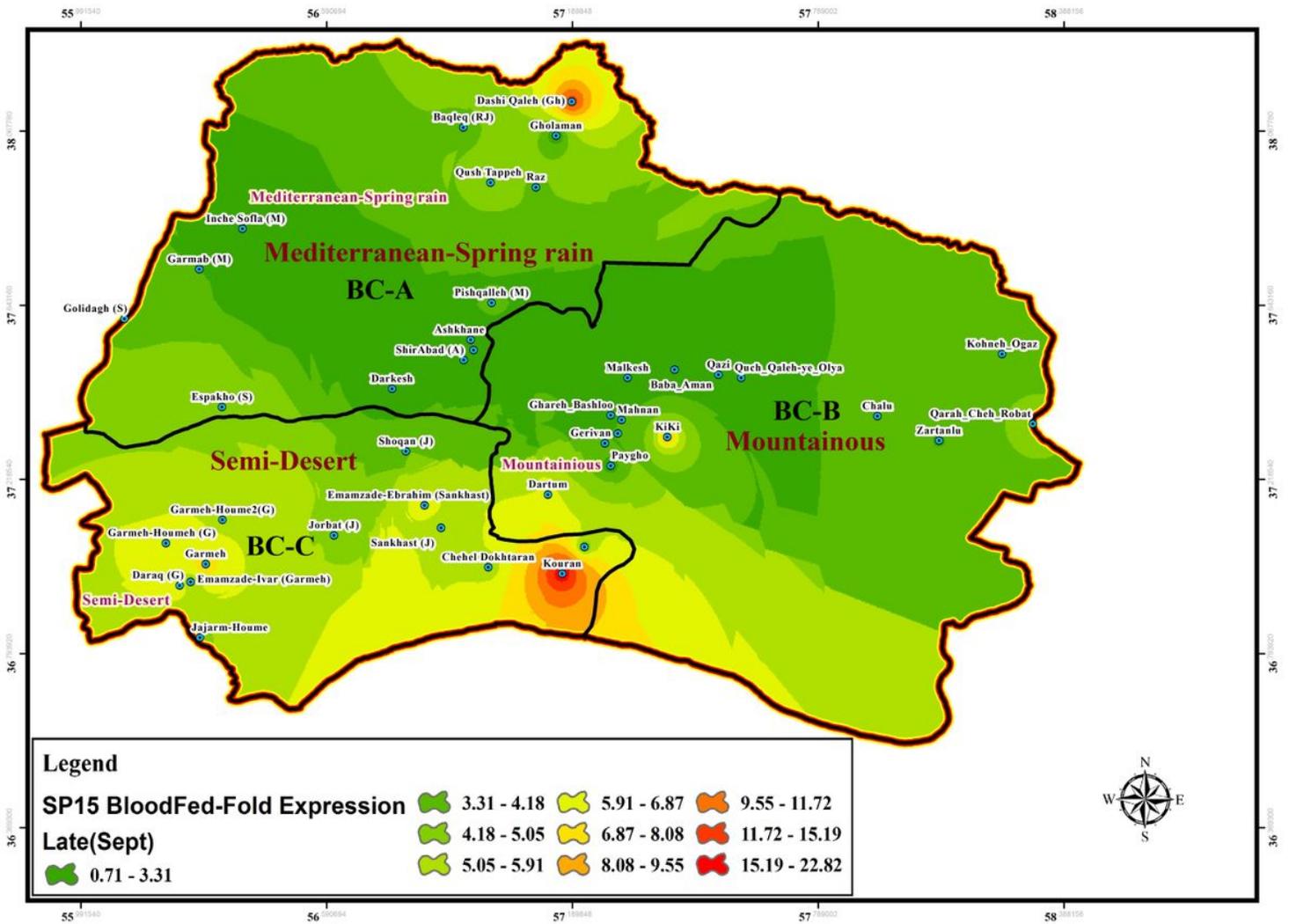


Figure 8
 Gene expression changes of SP15 salivary protein in wild-collected *Ph. papatasi* sand flies based on bio-climate zones and time of collection, September (Late), BC: Bio Climate regions.

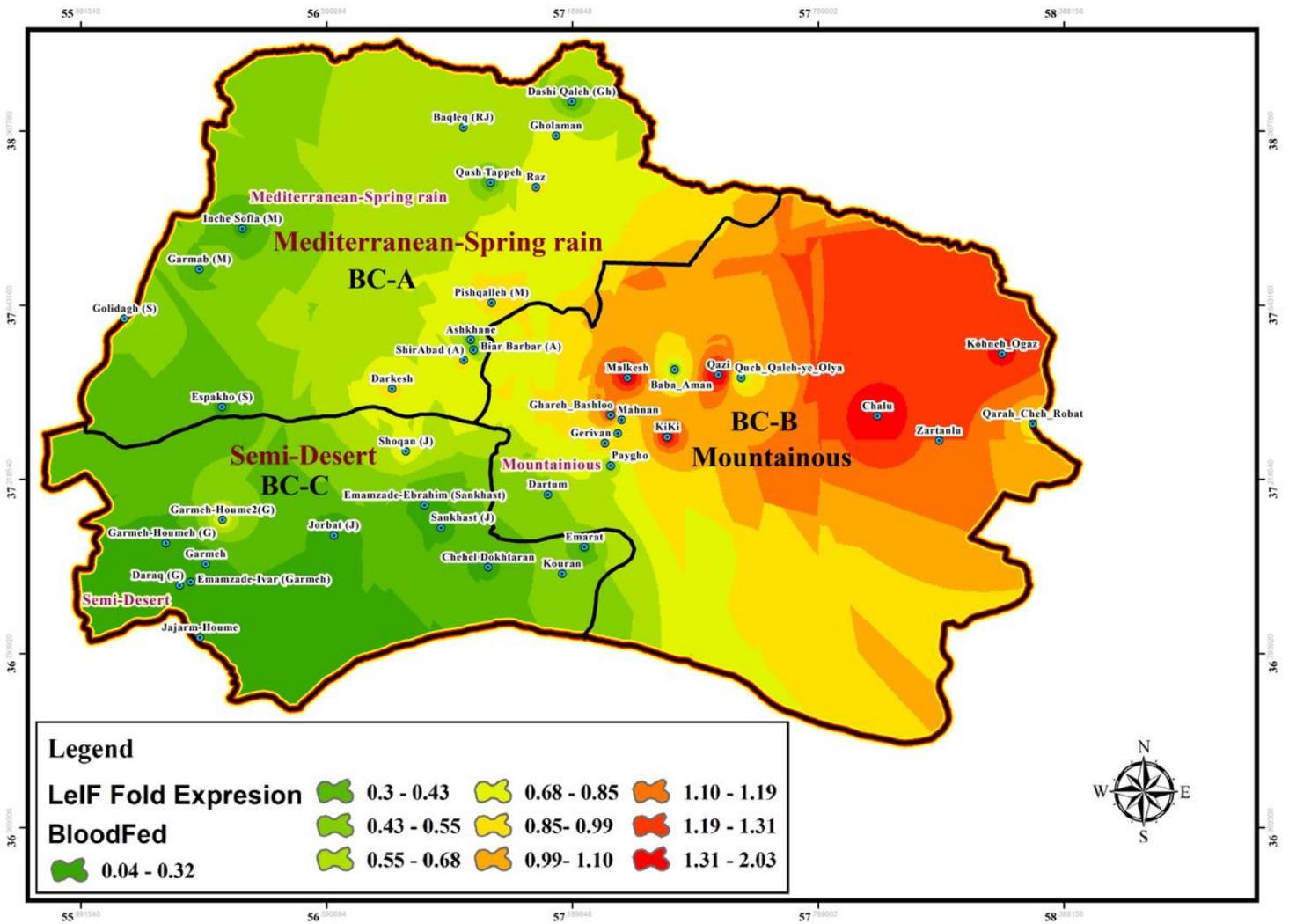


Figure 9

Fold change expression of Leishmania eukaryotic initiation factor (LeIF) gene based on "Blood-fed" physiological status of natural-collected *Ph. papatasi* sand flies in three bio-climate regions of studied area during June-September, BC: Bio Climate sites.

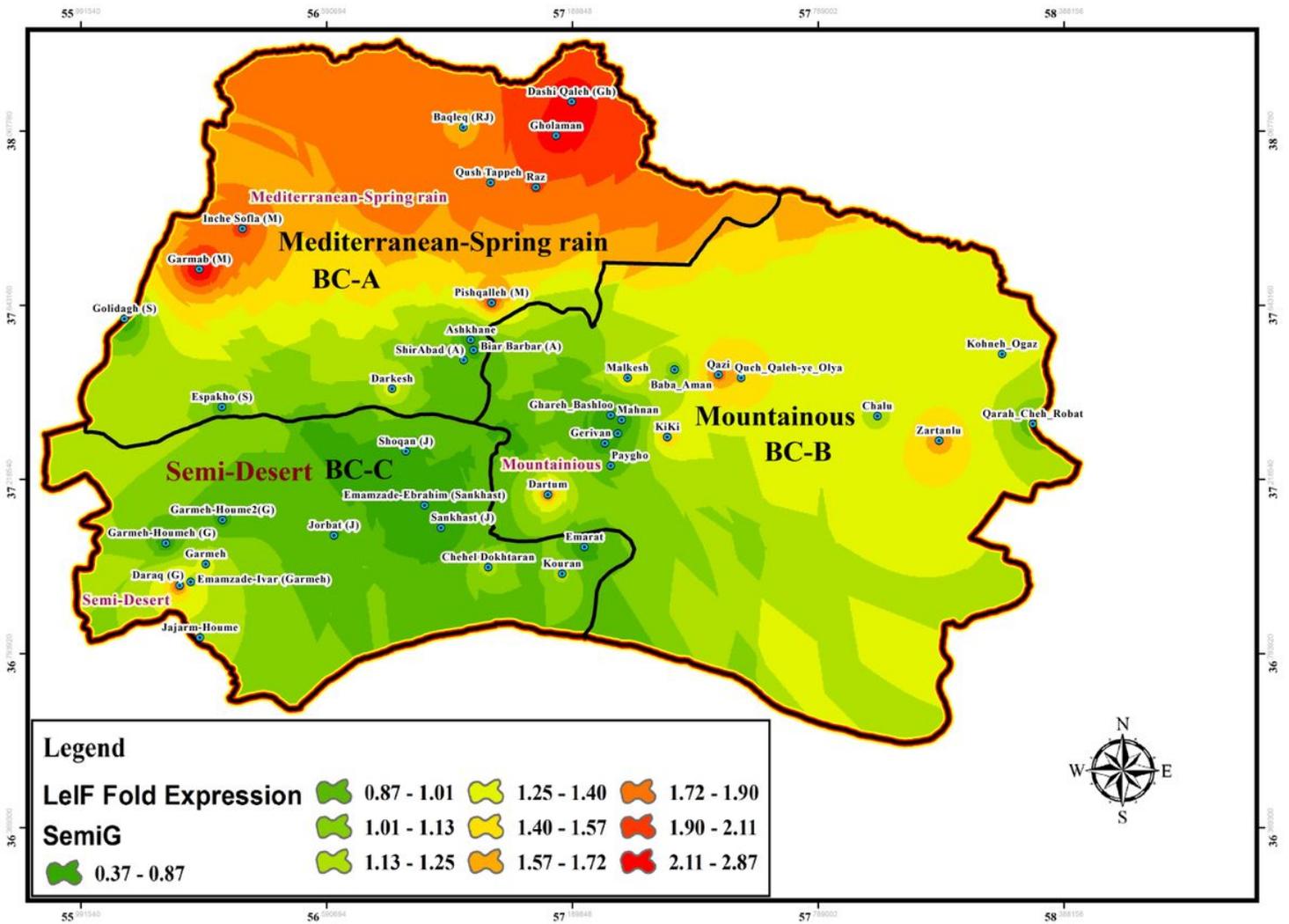


Figure 10

Fold change expression of LeIF gene based on "Semi-gravid" physiological status of wild-caught *Ph. papatasi* population in three bio-climate regions of studied area during June-September, BC: Bio Climate locations.

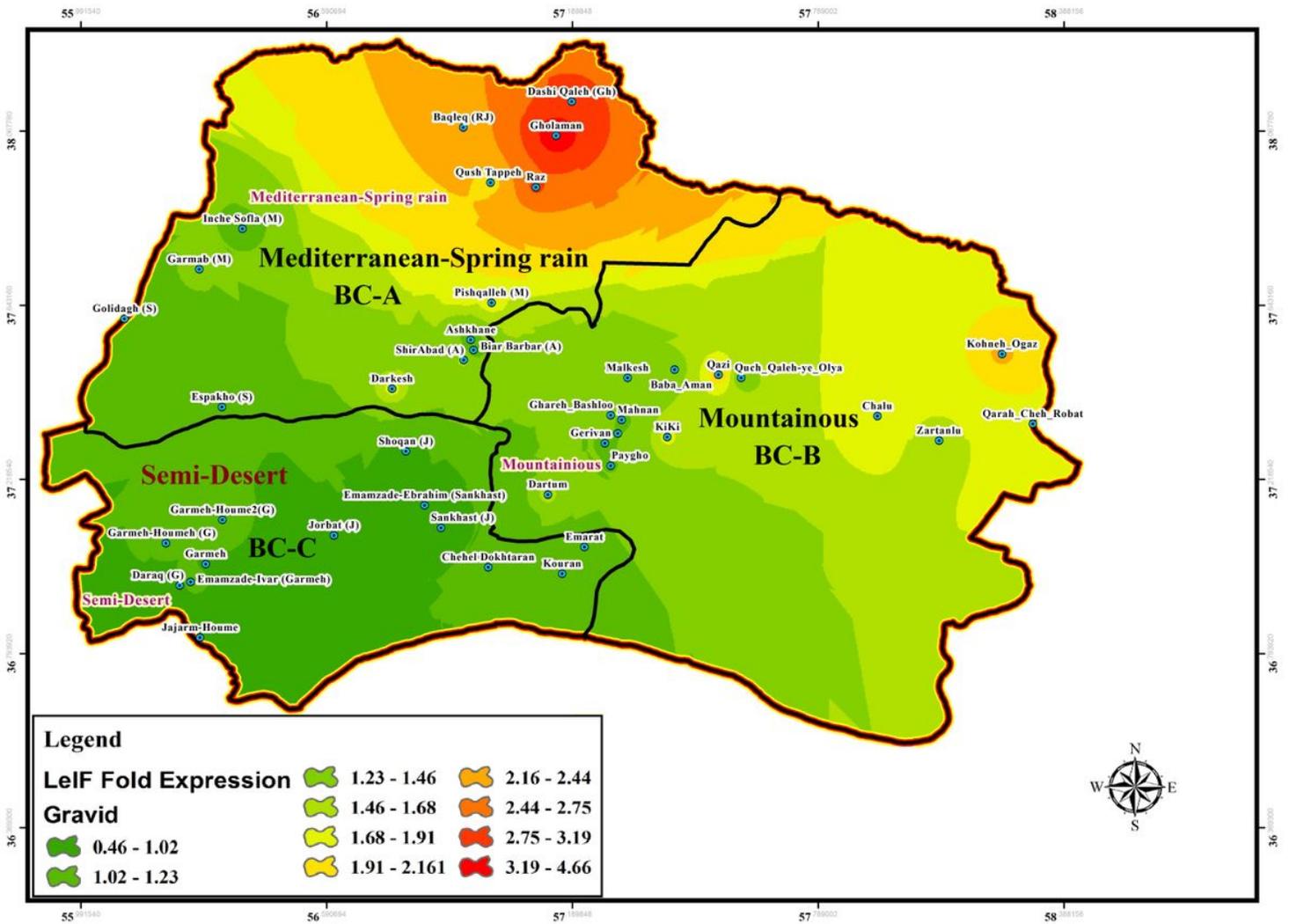


Figure 11

Fold change expression of LeIF gene based on "Gravid" physiological status of field-collected *Ph. papatasi* population in three bio-climate regions of studied area during June-September, BC: Bio Climate regionalized parts.

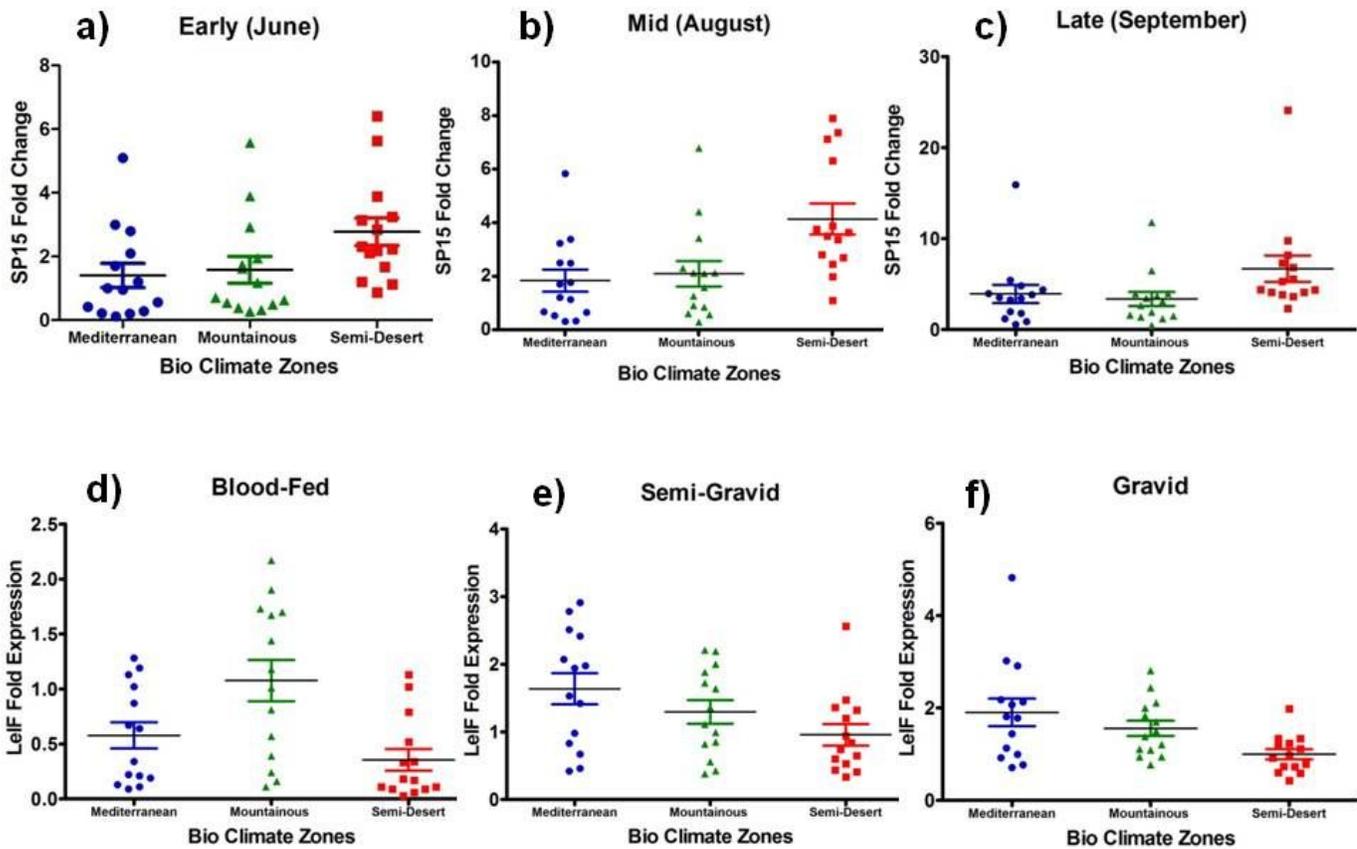


Figure 12

Salivary gland (SP15) and Leishmania (LeIF) genes expression from field-collected *Ph. papatasi* at northeast of Iran. Genes expression profiles were evaluated as fold changes (Y axis, SP15 and LeIF) over the calibrator, using the $2^{-\Delta\Delta CT}$ method. Expression profiles of *P. papatasi* collected in June (early in the season), August (middle), and September (late) 2017-2018 are displayed for SP15 (a, b, c), and also for LeIF with different physiological statuses, a) Blood-fed, b) Semi-gravid, c) Gravid. Horizontal bars represent the expression mean values between the samples and each sample represents an individual fly. Circle, triangle and square represent each related bio-climate zones: Mediterranean, Mountainous, and Semi-desert, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ADDITIONALFILES.docx](#)
- [Additionalfile4TableS2.xlsx](#)
- [Additionalfile7TableS5.docx](#)
- [Additionalfile6TableS4.txt](#)
- [Formula1.jpg](#)
- [Additionalfile5TableS3.xlsx](#)
- [GraphicalAbstract.png](#)